# Groundedness

# GROUNDEDNESS

JÖNNE KRIENER

Its Logic and Metaphysics

PhD
Philosophy
Birkbeck College, University of London

# DECLARATION

The work presented in this thesis is my own.

*Palaiseau, March 13th 2014*

Jönne Kriener

# ABSTRACT

In philosophical logic, a certain family of model constructions has received particular attention. Prominent examples are the *cumulative hierarchy* of well-founded sets, and Kripke's least fixed point models of *grounded* truth. I develop a general formal theory of groundedness and explain how the well-founded sets, Cantor's extended number-sequence and Kripke's concepts of semantic groundedness are all instances of the general concept, and how the general framework illuminates these cases. Then, I develop a new approach to a grounded theory of proper classes.

However, the general concept of groundedness does not account for the philosophical significance of its paradigm instances. Instead, I argue, the philosophical content of the cumulative hierarchy of sets is best understood in terms of a primitive notion of ontological priority.

Then, I develop an analogous account of Kripke's models. I show that they exemplify the *in-virtue-of* relation much discussed in contemporary metaphysics, and thus are philosophically significant. I defend my proposal against a challenge from Kripke's "ghost of the hierarchy".

# CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# PUBLICATIONS

Chapter 4 has appeared as 'The Groundedness Approach to Class Theory', 2014, *Inquiry: An Interdisciplinary Journal of Philosophy*. Section 7.2 is to appear in A. Malpass (ed.) *An Introduction to the History of Philosophical and Formal Logic: From Aristotle to Tarski*, contracted with Continuum Press.

# ACKNOWLEDGMENTS

This work was supported by the *European Research Council* and the University of Oslo. I thank the Institut Jean Nicod, Paris, for their hospitality.

I am indebted to my supervisor Øystein Linnebo. His constructive criticism has shaped this material significantly.

*Chapters 2 and 8*: I thank Jon Litland for many sharp questions, constructive criticism and inspiring discussions.

*Chapter 4*: I thank Casper Storm Hansen, Sam Roberts and participants at seminars in Bristol and Oslo.

*Chapters 5 and 6*: I am indebted to Michael Potter for turning my attention to this way of arriving at a metaphysical reading of groundedness.

*Section 7.2*: I thank Antje Rumberg for patiently sharing her extensive knowledge of Bolzano's theory of logic.

*Chapter 9*: I am very grateful to Albert Visser and Philip Welch for corrections on earlier drafts as well as many inspiring questions.

# 1 INTRODUCTION

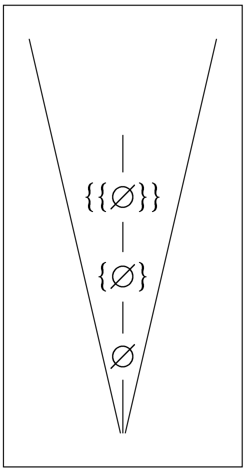


Figure 1: The Cumulative Hierarchy of Sets

The subject of the present study are certain structures that have received particular attention in mathematically informed philosophy. Prominent among them are the *cumulative hierarchy* of well-founded sets, and Kripke's least fixed point models of *grounded* truth. For present purposes, it suffices to sketch these well-known paradigms in broad strokes. Details will be filled in later.

The well-founded sets are those which can be generated from nothing, by iterated set formation. They are usually pictured as in figure 1. Kripke's least fixed point contains all and only those truths in a language with truth predicate that can be generated from true sentences without truth predicate, by truth introduction and closure under some monotone logic. We can also give a picture of these grounded truths, see figure 2.

In the present study, I will seek to answer two broad questions. Firstly, what is it that these cases have in common? I will develop a general concept of *groundedness* that captures both paradigms and

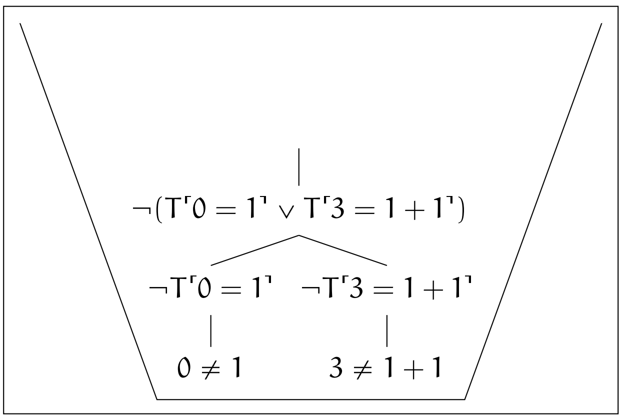


Figure 2: Grounded Truth

illuminates their connection. Secondly, as philosophers, why are we interested in them? I will argue that the philosophical significance of groundedness needs to be explained, and will develop and defend one such account.

The analogy between grounded truths and well-founded sets has occasionally been noted, but is obscured by the way in which Kripke's model construction is usually presented. To overcome this difficulty I develop a general theory of groundedness. Its primitive is the notion of a generator, a way of generating something from some things, broadly construed. Intuitively, it takes some things and produces something, possibly among them, possibly not. Several things may thus be produced from the same material, just as the same thing may be generated from distinct pluralities. Formally, a generator behaves like a many-to-one relation.

In a nutshell, I say that something is grounded in some things if it is generated from them, or obtained by iterated generation. I show how this general concept of groundedness illuminates the case of grounded truth. For example, it clarifies that two distinct ways of generating truths feed into semantic groundedness. Firstly, we generate truths of the form 'It is true that $\phi$' from the truth that $\phi$. Secondly, we generate syntactically complex truths by closure under logic. It is only this latter generation that varies across the range of Kripke fixed point constructions. Thus, my presentation renders precise the common core of the various constructions: the Kripke truth generator.

Then, I focus on one case of potentially great philosophical relevance: groundedness models for type-free theories of concept-extensions. Unlike the well-foundedness of sets and Kripke's models of truth, this instance of groundedness has not yet been sufficiently developed. I present new methods for obtaining theories of grounded classes, and test them against antecedently motivated desiderata. My findings cast doubt on whether a theory of grounded classes can accommodate both the extensionality of classes and allow for class definition in terms of identity.

I then change the perspective from an interest into the logical properties of groundedness, to asking for its philosophical significance (ch. 5). Thus, I move on to my second overall question. Here, I find that much work is yet to be done. I argue that the general concept of groundedness does not account for the philosophical significance of its paradigm instances. This problem has to my knowledge not yet been sufficiently appreciated in the literature on semantic groundedness. One contribution of my thesis is to clearly formulate this challenge.

I engage critically with Forster's [2008] case that the Church-Oswald model of class theory is as legitimate as the cumulative hierarchy of sets (§ 5.2). Forster's highly original paper has unfortunately not received much attention, and I inted to fill this gap. I carry out a

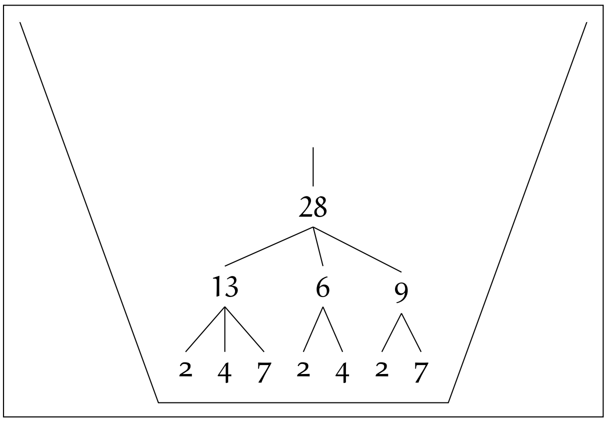


Figure 3: Insignificant groundedness: The *Sum* generator

thorough investigation and conclude that his argument is inconclusive. Then, I give other examples of groundedness that arguably lack philosophical significance. For instance, any plurality of natural numbers may be viewed as generating their sum. Then, we may call 28 *sum*-grounded in 2,4 and 7: generate 13 from them together, 6 from 2 and 4, and 9 from 2 and 7, and then obtain 28 as the sum of 13, 6 and 9 (see figure 3). In fact, there is an inexhaustible stock of such trivial groundedness.

From this discussion I conclude that we need to look elsewhere and supplement the formal concept with philosophical content. In the second half of my thesis, I will take first steps into this direction. Doing so, I initiate a novel approach to groundedness and connect areas of logical-philosophical research that, with few very recent exceptions, have been separate so far.

My starting point will be a certain body of literature in the philosophy of set theory (chapter 6). More precisely, in order to supplement groundedness with philosophical content I concentrate on the well-foundedness of sets, and engage with views as to its philosophical significance. I focus on discussions as to how the *iterative conception* of sets is to be explicated. Usually, the philosophical core of the iterative conception is glossed by saying that elements are prior to their set, or that a set is constituted from its elements. Authors disagree about what to make of this priority. From key texts of the past four decades I extract the following stance: the priority of elements over their set cannot be that it is constructed from them, nor that the set could not exist without them – unless we understand the latter in terms of a *sui generis* modality. Instead, the relation between elements and their set exemplifies a basic philosophical notion of constitution (§ 6.3). This notion of constitution is characterized by examples and structural principles such as non-circularity. To this extent we understand constitution, even if we have not defined it. Moreover, the notion thus characterized is philosophically significant. At least, this is an assumption central to my study. It is rendered plausible by the wide interest among philosophers into constitution and closely related notions.

On this basis, I propose the following partial answer to the challenge from chapter 5: Even if groundedness in general is not guaranteed to be philosophically significant, its paradigm case of the well-founded sets is, because the set generator exemplifies the philosophical notion of constitution, or ontological dependence. The relation between a set and its elements is frequently given as an example of ontological priority. In fact, proposed analyses of ontological dependence are tested against the relation between a set and its elements. In this precise sense, set generation does philosophical work: it guides research into ontological priority.

I then return to the other paradigmatic case of groundedness, Kripkes' least fixed point models of truth, and ask for a way of supplementing it with philosophical content. At this point, my general concept of groundedness pays off. Since both paradigms have been brought into the same general form and we know how to account for the well-founded sets, my general framework suggests an analogous case for semantic groundedness. In the remainder of the thesis I develop and defend the following view: Kripke's models are philosophically significant because Kripke's generator exemplifies the philosophical notion of a truth holding in virtue of others, in the same sense that the set generator tracks the idea that a set is constituted from its elements. In a slogan, the in-virtue-of relation is for Kripke's least fixed point what constitution is for the cumulative hierarchy of sets. To the best of my knowledge, the present study is the first to take this approach to semantic groundedness.

I begin to develop this view by explaining the relevant notion of *truth in virtue of* (ch. 7). My presentation follows its venerable history. I lay out Bernard Bolzano's theory of the in-virtue-of relation (§7.2). Remarkably, he characterizes *Abfolge* in the same way as philosophers of set theory have characterized the priority of elements over their set. Bolzano does not attempt to define the in-virtue-of relation, and expresses doubt that this can be done at all. However, he is confident that the notion is grasped from example that he gives, as well as by reflection on formal principles.

Then, I turn to recent work by Kit Fine who has revived much of Bolzano's approach, and has put the in-virtue-of relation to new use in contemporary metaphysics. This work culminates in Fine's [2012b] and its *pure logic of ground*. I settle on it as the regimentation of the philosophical notion of a truth holding in virtue of others.

Thus, I arrive at examples and formal principles that characterize the in-virtue-of relation. They do so in much the same way as the notion of constitution was characterized. In its terms, again, I was able to account for the philosophical significance of set groundedness.

I develop an analogous case for the significance of Kripke's least fixed point constructions (ch. 7): I propose to view semantic groundedness as exemplifying the in-virtue-of relation.

On the one hand, I provide evidence from the literature that research into the in-virtue-of relation is guided by the thought that, say, it is true that snow is white in virtue of snow being white, much like research into ontological constitution is guided by the example of a set being constituted from its elements. In this sense, Kripke's truth generator captures an idea that does philosophical work.

On the other hand, I show that the connection between semantic groundedness and the in-virtue-of relation is robust, since the former satisfies the formal principles by which the latter is characterized. I do so in two steps, based on my analysis of semantic groundedness in terms of a truth generator **T**, and a logic generator **W**. Firstly, I show that Fine's *pure logic of ground* together with axioms 'It is true that $\phi$ because $\phi$' and 'It is not true that $\phi$ because $\neg\phi$', is complete with respect to **T**-generation and its transitive closure. Secondly, I present a set of rules that, according to Fine and other leading researchers, characterize how the in-virtue-of relation interacts with logic. I add these to the *pure logic of ground*, and show that the resulting system is sound, and partly complete, with respect to **T**-**W**-generation and its transitive closure.

The cumulative hierarchy provides a formal model of those principles which have been assumed for constitution. I show that **T**-**W**-groundedness models structural principles for in-virtue-of together with axioms about in virtue of what something is true. In this precise sense, semantic groundedness stands to the in-virtue-of relation as set groundedness stands to constitution. This is my account of the philosophical significance of Kripke's least fixed point models.

In the final chapter of my thesis I consider and answer an objection. If my case for the philosophical significance of semantic groundedness can only be made in a meta-language, the objection goes, I cannot account for the groundedness of truth in our own language. I explain that this challenge amounts a variant of Kripke's own complaint that "... the ghost of the hierarchy is still with us" (§ 9.2), and I develop a novel response to it. Briefly put, I use the intensional logic of *well-ordered* time to enable an axiomatic theory of truth to characterize the stages of Kripke's construction, and thus the formal concept of semantic groundedness.

# 2

# A GENERAL THEORY OF GROUNDEDNESS

## 2.1 INTRODUCTION

Groundedness has figured prominently in the literature on the semantic paradoxes.[1] In this chapter, I will develop a generalized concept of groundedness. It does not only apply to sentences, propositions, or to the truth-values of sentences or propositions. Whenever we are given some things we may ask whether, how, and in what they are grounded.

What does it mean, at this general level, for something to be grounded? Recently, Thomas Forster has proposed a generalized iterative conception of sets [2008].[2] My goal in this chapter is to develop his idea into a general theory of groundedness. It not only subsumes existent accounts, but also illuminates how they are connected.

Forster's generalized iterative conception is closely linked to what he calls "*recursive datatypes*" [2008, p. 99]. Outside of computer science these are better known as *inductive definitions*, and have been examined thoroughly in the 1970s, most prominently by Yannis Moschovakis [1974]. In particular, I will make use of tools from Peter Aczel's 1977 handbook entry, a source that is regrettably little used in the philosophical literature on groundedness. One goal of this chapter is to make available insights of Aczel's for the discussion of groundedness in philosophical logic.

Mathematically, this chapter's general theory of groundedness will be included in the theory of inductive definitions. Philosophically, there is more to groundedness. However, developing the philosophical side of the concept I postpone to later chapters, beginning in chapter 5.

One intended application of the following is to the universe of sets itself. In particular, I will show that they are all grounded (§ 2.7). Of course, however, there is not set of all sets. Therefore, I will develop my general concept of groundedness not by the standard set theoretic means of mathematics. Instead, I will work within *plural* logic.[3] For its primitive of a thing *being among* some things I adopt Burgess' notation '$\propto$'.[4]

---

1 Herzberger [1970]; Kripke [1975]; Yablo [1982]; McCarthy [1988]; Maudlin [2004]; Leitgeb [2005]

2 Forster uses his generalized iterative conception to argue for the legitimacy of certain non-standard set theories. As I will explain in chapter 5 below, I do not think his argument is conclusive.

3 In this respect, too, I go beyond Forster's proposal. However, I emphasize that my choice of a plural logic framework is for merely practical reasons, and nothing hinges on it. An other framework would do, too, as long as it allows for foundations of set theory. One such alternative framework is higher-order logic on a Fregean interpretation, another may be category theory.

4 Burgess motivates this choice as follows.

> Much as the symbol used in set theory for 'element' is a stylized epsilon '$\in$', the symbol used here for 'is among' is a stylized alpha '$\propto$'. [Burgess, 2004, p. 197]

For simplicity, I will use the singular locution 'plurality' to refer to some things.[5] I will also use '$xx \sqsubseteq yy$' as short for '$\forall z(z \propto xx \to z \propto yy)$'.[6] Further, I will assume that my plural metalanguage has a plural term forming operator that I denote by the comma sign. Thus, $x, y, z$ is a plural term, as is $xx, y$. Of course, the comma will also keep its usual syntactic role; the ambiguity will always be resolved in context. Finally, I will use three dots '...' (read: "and so on") as a natural way of denoting infinite pluralties: $0, 1, \ldots$. This general plural meta language will prove particularly useful when spelling out the groundedness of ordinals (§ 2.6) and sets (§ 2.7).

## 2.2 GENERATORS

Forster formulates his iterative conception in terms of *constructor* functions. I will not adopt this terminology. In my study, issues from the philosophy of mathematics will play an important role. In this context, 'constructor' is not a sufficiently neutral word. Therefore, I use the term 'generator' which I hope not to provoke those philosophical associations. It is certainly intended as a neutral label within a general framework.[7]

The concept of a *generator* is the primitive of my theory of groundedness. For intuition, think of a generator as a recipe by which something is obtained from some things. Formally, a generator $\gimel$ is a relation whose first argument place takes plural and whose second argument place takes singular terms.[8] Thus, generation statements are of the form

$$yy \gimel x$$

where, however, $yy$ may be one thing or indeed none.[9] Note that this general concept of a generator allows for cases in which two distinct things are generated from the same things.

Using '$\propto$', I deviate from the mainstream that uses '$\prec$' to denote the relation of something being among some others. My reason for deviating is that I will have to use the symbol '$\prec$' for another notion (definition 5). At any rate, doing so I am not committed to any claim about plural logic as a suitable regimentation of natural language plural locutions.

5 However, it is important to keep in mind that this term can always be paraphrased away in plural terms.

6 As here, I will frequently be lax about the use-mention distinction. For example, I will use simple quotation marks where Quine corners would be more accurate, but less readable.

7 Kit Fine [1991] uses the label 'constructor' for a framework that to my understanding has also a broad range of applications. In fact, Fine's first example of *constructional ontology* is the cumulative hierarchy of sets, that is also one focus of the present study.
However, unlike much material of later chapters (7,8), the present general theory of groundedness does not build on work of Fine's.

8 The Hebrew letter gimel '$\gimel$', pronounced much like the initial sound of 'groundedness', is a variable for generators, as are subsequent letters of the Hebrew alphabet.

9 Thus, I allow for what one may call *empty pluralities*.

Examples abound. The *formation rules* of a formal language describe a generator: a disjunction $\phi \vee \psi$ is generated from a formulae $\phi$ and $\psi$, as is a conjunction $\phi \wedge \psi$. Other examples are the introduction rules of a formal proof system. In propositional logic, a theorem $\phi \vee \psi$ is generated from theorems $\vee$.

Forster's constructors are functions. I lift this restriction and work with relations. Thus, I do not need what Forster calls "destructors", functions from something to those things from which it is generated. Instead, I can take the generator's inverse. Further, unlike Forster I consider the ways in which we generate objects from pluralities. Let me give an example. Consider the language of propositional logic based on propositional letters $p, q, r$, with decorations. From these atomic sentences we *generate* complex sentences using the following formation rules:

$$\begin{array}{cc} P_1 \; \dfrac{\phi}{(\neg\phi)} & \dfrac{\phi \qquad \psi}{(\phi \vee \psi)} \; P_2 \\[2ex] P_3 \; \dfrac{\phi \qquad \psi}{(\phi \wedge \psi)} & \dfrac{\phi \qquad \psi}{(\phi \rightarrow \psi)} \; P_4 \end{array} \tag{1}$$

Presently, what matters are the rules $P_1$ to $P_4$ taken together. Together, they describe one generator **P**. This generalizes Forster's terminology, who would speak of four constructors $P_1$ to $P_4$.

In case that the things at hand form a set, I will often use set-theoretic resources to speak of them and how they are generated. This will render the presentation more familiar. For example, as the sentences of propositional logic form a set, I represent the generator $P$ as the union of the following relations.

$$P_1 \coloneqq \{\langle X, \zeta \rangle : X = \{\phi\}, \zeta = (\neg\phi)\} \tag{2}$$
$$P_2 \coloneqq \{\langle X, \zeta \rangle : X = \{\phi, \psi\}, \zeta = (\phi \wedge \psi)\} \tag{3}$$
$$P_3 \coloneqq \{\langle X, \zeta \rangle : X = \{\phi, \psi\}, \zeta = (\phi \vee \psi)\} \tag{4}$$
$$P_4 \coloneqq \{\langle X, \zeta \rangle : X = \{\phi, \psi\}, \zeta = (\phi \rightarrow \psi)\} \tag{5}$$

However, I do not intend to reduce the notion of a generator to that of a relation understood set-theoretically, plurally or by some other means. A generator $\gimel$ is a way of obtaining an $x$ from some $yy$, a rule how to move from $yy$ to $x$. It is not a collection of pairs $\langle yy, x \rangle$.[10] Rather, $\gimel$ is the intension corresponding to such an extensional characterization.

A generator $\gimel$ is all that is needed to formulate my general concept of groundedness. I will give two ways of characterizing some $y$ as $\gimel$-grounded in $xx$. They are equivalent, but formalize intuitively distinct ideas. Hence, it will prove useful to have available both the one and the other characterization of groundedness.

10 Officially, '$\langle yy, x \rangle$' is short for the pairs $\langle y, x \rangle$ such that $y$ is among $yy$. It is compatible with my use of plural logic to understand these pairs in the standard set-theoretic way due to Kuratowski.

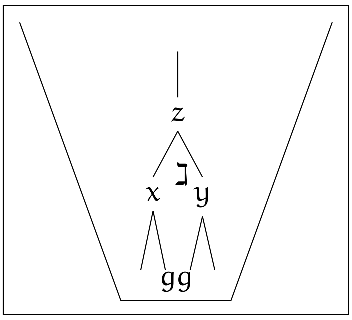


Figure 4: $\gimel$-Groundedness in $gg$

Given a generator $\gimel$, we can ask two prima facie distinct questions. Firstly, we can ask what may be generated from some things. Secondly, we can ask what something is generated from. Answering the first question we characterize groundedness in terms of as what can be generated. Indulging in spatial metaphor we may call this an *upwards* characterization of groundedness. Formally, the first definition will identify $y$ as grounded in $xx$ if it is arrived at by iterated $\gimel$-generation, starting from $xx$.

Answers to the second question characterize groundedness by tracing *downwards* what something is generated from. This characterization of groundedness motivates an alternative definition according to which $y$ is called grounded in $xx$ if tracing down what $y$ is generated from, we end with $xx$.

To the best of my knowledge, the first author to spell out these two characterizations of groundedness in a philosophical context was Stephen Yablo [1982]. Much of the following may be viewed as a elaboration on Yablo's general, formal theory of groundedness. I will state and explain these connections as they arise.

## 2.3 UPWARDS: GENERATION

Assume we are given some $gg$. Then, let us call $x$ $\gimel$-grounded in $gg$ if $x$ is among $gg$, if it stands in the relation $\gimel$ to ("is generated from") some $yy \sqsubseteq gg$, or if $x$ is generated from some $zz$ each of which is already grounded in $gg$. This idea is visualized well by drawing, as in figure 4, a funnel-shaped diagram whose base represents $xx$, and every point in its area represents something grounded in them.

To render precise this idea in the present, very general context, I need to clarify what it means to *iterate* generation. Given a well-ordering, we can define the *stages* of an iterated $\gimel$-generation along the well-ordering. Fortunately, for some things $ww$ to well-order some other things $yy$ can be expressed in our present, plural setting as a suitable plurality of pairs [Shapiro, 1991, p. 106]. For readability, if $ww$ is a well-ordering of $yy$, I will use expressions of the form '$w+1$'

for the *ww*-successor of *w*. In cases where $yy$ form a set I will, for simplicity, work with its order-type, the ordinal $\alpha$.

Assume that *ww* are well-ordered. I wish to formalize the iteration of $\gimel$ along *ww*, starting from $gg$. Let $0_{ww}$ be the least. Then, encode the first stage of our iteration of $\gimel$ as pairs $\langle 0_{ww}, g\rangle$ where $g$ is among $gg$. Given some $w$ that is among *ww*, the $w+1$th stage of our iteration of $\gimel$ is encoded by pairs $\langle w, x\rangle$, where $x$ is among the $w$th stage, or there are some $xx$ among these and $xx\gimel x$. Generalizing standard notation from the theory of inductive definitions, I will denote the things at the $w$th stage, the $x$ for which there is a pair $\langle w, x\rangle$, by '$I^w_\gimel(gg)$'. For $w$ limit among *ww*, let $I^w_\gimel(gg)$ be all the $I^{w'}_\gimel(gg)$, for $w'$ *ww*-preceding $w$, taken together.

**Definition 1** (Groundedness)**.** Let $\gimel$ be a generator and let $gg$ be some things.

$x$ is grounded in $gg$ by $\gimel$ ('$\gimel$-grounded in $gg$', in symbols: $gg \leqslant_\gimel x$) iff there is some well-ordered *ww* and some $w$ among *ww* and $x$ is among $I^w_\gimel(gg)$.

It will be useful to speak of this $w$ as the $\gimel$-$gg$-*rank* of $x$ with respect to a well-ordering *ww*.

Let us call $x$ *strictly* $\gimel$-grounded in $gg$ ('$gg <_\gimel x$') iff $gg \leqslant_\gimel x$ but $x$ is not one of $gg$.

$x$ is strictly $\gimel$-grounded in $gg$ if its rank is greater than $0_{ww}$. Intuitively, it is strictly $\gimel$ grounded in $gg$ if it takes at least one step to generate $x$.

Note that, in addition to what is grounded, groundedness involves three parties. Firstly, we can speak of something as grounded only in the sense of being grounded in some specific way, which is captured by a generator $\gimel$. Groundedness is $\gimel$-groundedness. Secondly, $\gimel$-groundedness is relative to some things $gg$ from which we start to $\gimel$-generate things, and arrive at what is $\gimel$-grounded in them. Groundedness is $\gimel$-groundedness in some $gg$.

Finally, note that for there to be anything grounded, it must be among things well-ordered by *ww*. They are used to render precise that $\gimel$-groundedness in $gg$ is being generated from $gg$ by *iterating* $\gimel$. Each of *ww* functions as one step in this iteration. Groundedness is $\gimel$-groundedness in some $gg$ along some well-ordered *ww*.

However, what is grounded need not to be one of *ww*. We may take seriously the label 'generation' and view an object $x$ grounded in $gg$ as having been created from them. Some cases of groundedness that I will consider later invite such a reading (sections 2.6, 2.7). Others, however, do not (p. 28 below, and chapter 3), and the ontological reading is not part of my general framework.

A mundane example will show just how common groundedness is. Recall the generator **P** of the previous section. The sentences of propositional logic are **P**-grounded in the propositional letters. For every sentence $\phi$ of propositional logic there is a well-ordering such that

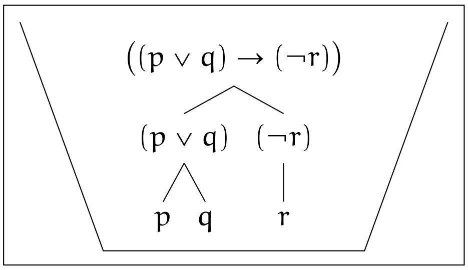


Figure 5: The Grounding Cone of Propositional Logic

$\phi$ is generated from some propositional letters by **P**-generation along the well-ordering (see figure 5). Thus, $(p \vee (q \wedge r))$ is **P**-grounded in $p, q$ and $r$.

Note that $x$ is trivially $\gimel$-grounded in $gg$ if it is one of them. If $x$ is one of $gg$ then $x$ is $\gimel$-grounded in $gg$ without being generated but, so to speak, as part of the ground.

Such groundedness is to be distinguished from cases where $x$ is $\gimel$-generated, but $\gimel$ allows us generate something from nothing.[11] The most prominent case of having been generated from nothing is the empty set (see §2.7 below). Another example is the generation of a class from those things that are not its members, as discussed by Forster [2008], that allows for the generation of the universal class from nothing (see §5.2). Generally, in such cases of generation from nothing it makes sense to speak of groundedness in nothing: $y$ is $\gimel$-grounded in nothing iff it is generated from nothing, by iterated application of $\gimel$ along some well-ordering. Goundedness in nothing, is of course not ungroundedness. Ungroundedness is just not to be grounded, and thus like groundedness a tertiary notion: for something to be ungrounded with respect to a generator $\gimel$ and $gg$ is simply for it not to be $\gimel$-grounded in $gg$. $gg$ may be nothing, in which case something is $\gimel$-ungrounded in nothing if it cannot be $\gimel$-generated from nothing, directly or indirectly. For example, $(p \vee q)$ is not **P**-grounded in nothing.

If we have two generators $\gimel$ and $\daleth$ they can be combined and give rise to a more inclusive notion of $\gimel$-$\daleth$-groundedness. This is done as follows. Let $gg$ be some things and let $\gimel$ and $\daleth$ be generators. Now we may obtain things from $gg$ either by means of $\gimel$ or by means of $\daleth$: but this is a new way of generating things, a combined generator $\gimel$-$\daleth$. Think of $\gimel$ as the rule to infer $y$ if $xx$ are so and so, and of $\daleth$ as the rule to infer $y$ if $xx$ are such and such. The combined generator $\gimel$-$\daleth$ then is the rule which allows us to infer $y$ from $xx$ if they are so and so *or* such and such. Thus, $xx\gimel$-$\daleth y$ iff $xx\gimel y$ or $xx\daleth y$.

For example, let a generator **M** be given by the following rules.

11 The difference between not being generated, and being generated from nothing is noted in [Fine, 2012b, p. 47]. Fine relates it to a distinction relevant to the notion of one truth holding *in virtue of* another to which I will return in chapter 7.

$$M_1 \, \frac{\phi}{\Diamond\phi} \qquad \frac{\phi}{\Box\phi} \, M_2$$

The sentences of propositional *modal* logic then are **P**-**M**-grounded in the propositional letters.

We can define, relative to $\daleth$, an operator $\Gamma_\daleth$ that takes some things $xx$ and outputs exactly those $yy$ each of which is $\daleth$-generated from some $zz$ among $xx$. Formally,

**Definition 2.**

$$y \propto \Gamma_\daleth(xx) :\Leftrightarrow \exists zz \sqsubseteq xx(zz \daleth y)$$

This operator $\Gamma_\daleth$ allows for the following, useful re-characterization of $\daleth$-groundedness. Some $x$ is $\daleth$-grounded in $gg$, we may say, if starting from $gg$, and iterating this operator $\Gamma_\daleth$, we eventually find $x$ in its output. This is equivalent to saying that $x$ is $\daleth$-grounded in $gg$ iff $x$ is in the least plurality containing $gg$ and closed under $\Gamma_\daleth$, in other words the least fixed point of $\Gamma_\daleth$ that contains $gg$. This characterization renders clear that the present concept of groundedness is closely related to the theory of inductive definitions, as in Moschovakis [1974] (see also Barwise [1975]; Aczel [1977]). In fact, for domains that form a set, to be $\daleth$-grounded is to be in the inductive set $I_{\Gamma_\daleth}$, the least fixed point of $\Gamma_\daleth$.

Before in the next section I turn to the *downwards* characterization of groundedness, let me give another example of groundedness. It is closely connected to the mundane case of complex propositional formulae (p. 24) and as such both simple and uncontroversial. At the same time, however, it sets the stage for the next chapter.

The example is that of truth, more precisely Tarski's inductive definition of it. Consider, as above, firstly propositional logic. Let $\mathcal{V}$ be a valuation function that assigns truth or falsity to the propositional letters. As usual, we define the set $\{\phi : \mathcal{V} \models \phi\}$ following Tarski's compositional clauses.

Consider those truths that are either atomic or the negation of an atomic formula. That is, consider the true *literals*. Now, we can view each truth in $\mathcal{V}$ as grounded in the literals true in $\mathcal{V}$, by a generator **V** given as follows.

$$\begin{array}{cc} \dfrac{\phi}{(\phi \vee \psi)} & \dfrac{\neg\phi \qquad \neg\psi}{\neg(\psi \vee \phi)} \\ \dfrac{\phi}{\neg\neg\phi} & \dfrac{\psi}{(\phi \vee \psi)} \end{array} \tag{6}$$

We have that for every formula $\phi$ of propositional logic, $\mathcal{V} \models \phi$ just in case $\phi$ is **V**-grounded in the true literals.

Now let $\mathfrak{M}$ be a model of some first-order language $\mathcal{L}$. For simplicity, I assume that $\mathcal{L}$ has a constant for all and only the objects of $\mathfrak{M}$'s domain. An $\mathcal{L}$-sentence is true in $\mathfrak{M}$ if and only if it is grounded in the $\mathcal{L}$-literals true in $\mathfrak{M}$, by the generator $\mathbf{W}$,[12] which is given by the rules in 6 together with the following two rules for $\forall$, the only quantifier in the language $\mathcal{L}$.[13]

$$\frac{\psi(a) \qquad \psi(b) \qquad \dots}{\forall x(\psi(x))}\ a, b, \dots \text{ are exactly the } \mathcal{L}\text{-constants} \tag{7}$$

$$\frac{\neg\psi(a)}{\neg\forall x(\psi(x))}\ a \text{ is some such constant} \tag{8}$$

In other words, the set of first-order formulae true in a model is the least set containing the true literals and closed under the rules in 6 and 7. This means that we can view the complete theory of an $\mathcal{L}$-model $\mathfrak{M}$ as the sentences $\mathbf{W}$-grounded in the $\mathcal{L}$-literals true in $\mathfrak{M}$. This is a well known fact, presented for example in [McGee, 1991, p. 110]. For future reference, let me make it explicit.[14]

**Fact 1** (McGee 1990, example 5.5)**.** *Let $\mathcal{L}$ be any first-order language, and $\mathfrak{M}$ any $\mathcal{L}$-model; let $\mathcal{L}^m$ be the extension of $\mathcal{L}$ by a constant for every object in $\mathfrak{M}$'s domain. Now, for every complex $\mathcal{L}^m$-sentence $\phi$, $\mathfrak{M} \models \phi$ iff $\phi$ is $\mathbf{W}$-grounded in those $\mathcal{L}^m$-literals true in $\mathfrak{M}$.*

## 2.4 DOWNWARDS: PRIORITY

I now turn to develop a formal definition that captures the *downwards* idea of groundedness. Let $\daleth$ be a generator. We say that $x$ is immediately $\daleth$-prior to $y$ iff $x$ is among some things from which $y$ is $\daleth$-generated. For example, the propositional letters $p, q$ are each immediately $\mathbf{P}$-prior to their disjunction $(p \vee q)$. In general, immediate priority is not ensured to be irreflexive: some generators allow $y$ to be generated from some $xx$ such that $y$ itself is among $xx$. For example, from some natural numbers we may generate its least upper bound, which may be among them. We may view 7 as *least-upper-bound*-generated from 3,5 and 7, and as such prior to itself.[15]

*Mediate* priority is best developed via the following concept of an object's *priority tree* with respect to a generator. If we have some well-ordered things $ww$, we can encode ordered *sequences* $\langle x, y, \dots\rangle$ as some pairs $\langle w, x\rangle, \langle v, y\rangle, \dots$, where $v$ succeeds $w$ in the well-ordering. In this case, I will speak of $\langle x, y, \dots\rangle$ as a sequence *along* $ww$. Thus, we can also speak of one sequence extending, or being an initial segment of another.

12 '$\mathbf{W}$' for Tarski's 1935 German "*Wahrheit*". I reserve the letter 'T' for truth in the sense of Kripke (chapter 3).

13 Of course, the first rule reminds of the $\omega$-rule. Note, however, that the above is not intended to describe a formal system, but is thought of better as a way of constructing models.

14 A close kin of fact 1 will be proved as lemma 4 in section 3.5 below.

15 Of course, the generation of least upper bounds becomes more interesting once we consider infinite sets of numbers, see §2.6 below.

Some such sequences are a priority tree of an object $x$ if they track how $x$ is generated, in the precise sense of the following definition.[16]

**Definition 3** (Priority Trees)**.** Let $\daleth$ be a generator and $x, gg$ be some things. Let *ww* be some well-ordered things, and $\mathcal{T}$ some finite sequences along *ww*.

$\mathcal{T}$ are a $\daleth$ *priority* tree of $x$ along *ww* and ending in $gg$ (a '$\daleth$-$gg$-*ww*-tree' of $x$) iff

1. $x$ is the *root* of $\mathcal{T}$: there is exactly one sequence of length one among $\mathcal{T}$, and it is $\langle x \rangle$,
2. for every sequence among $\mathcal{T}$, every initial segment of it is among $\mathcal{T}$,
3. there is no infinite sequence of things $y_0, y_1, \ldots$ such that for all $n$, $\langle y_0, \ldots, y_n \rangle$ is among $\mathcal{T}$,
4. $\langle y, \ldots, z \rangle \sqsubseteq \mathcal{T}$ if and only if
   a) it has proper extensions among $\mathcal{T}$ and $z$ is $\daleth$-generated from all and only the $u$ such that $\langle y, \ldots, z, u \rangle$ are among $\mathcal{T}$, or
   b) $z \propto gg$, or
   c) $gg$ are nothing, and $z$ is $\daleth$-generated from nothing.

Some simple examples may help to parse the definition. $\langle \neg q \rangle, \langle \neg q, q \rangle$ are a $\mathbf{P}$-q-$1,2$-priority tree of $\neg q$ because the formula $\neg q$ is $\mathbf{P}$-generated in one step from the propositional letter $q$. $\langle (p \vee q) \rangle, \langle (p \vee q), q \rangle$ are not a $\mathbf{P}$-$p, q$-$1,2$-priority tree because it lacks $\langle (p \vee q), p \rangle$. Nor does $(p \vee q)$ have in $\langle (p \vee q) \rangle, \langle (p \vee q), q \rangle, \langle (p \vee q), p \rangle, \langle (p \vee q), r \rangle$ a $\mathbf{P}$-$p, q$-tree, this time because $r$ is not among the things that $(p \vee q)$ is $\mathbf{P}$-generated from.

I will frequently suppress mention of *ww*, $gg$, or even $\daleth$, if they are either irrelevant for the point at hand, or clear from context. Thus, it will be convenient to say that $\neg q$ has a $\mathbf{P}$-tree, or even just to say that it has a tree. Let us abstract slightly and speak of $y$ having a $\daleth$-priority tree *simpliciter* if for *some* $gg$ and *ww*, $y$ has a $\daleth$-$gg$-*ww*-priority tree.

To gain more intuitive access, we can characterize priority trees in graph-theoretic terms. For this, however, we need to assume that each object in our tree is represented by many distinct *tokens*. For example, recall that the sentences of propositional logic are $\mathbf{P}$-grounded in the propositional letters (p. 24). Figure 6 shows a $\mathbf{P}$-$\{p, q\}$-priority tree of $\big((p \wedge (q \rightarrow \neg p)) \rightarrow \neg q\big)$. This graph has two distinct vertices which are both labelled 'p'. In other words, these two vertices are both representations of the atomic formula $p$ that figures in both the generation of $\neg p$ and the generation of $p \wedge (q \rightarrow \neg p)$. For simplicity, however, I will frequently speak of priority trees graph-theoretically

16 Definition 3 transfers Aczel's 1977 definition 1.4.4 into the present, plural setting.

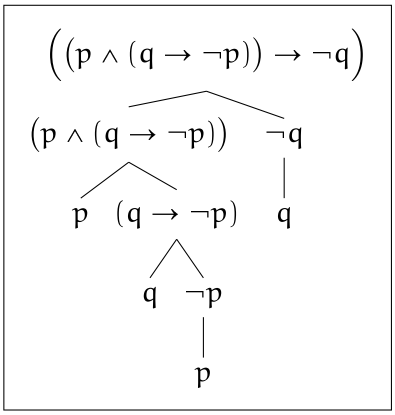


Figure 6: A $\mathbf{P}$,p, q priority tree of $\big((p \wedge (q \to \neg p)) \to \neg q\big)$

without explicitly distinguishing between things and their tokens. I can do so, because officially, a priority tree are just some sequences, in which something may occur more than once.

**Definition 4.** Let the *height* of a priority tree $\mathcal{T}$ be the least upper bound, relative to the well-ordered *ww*, of the heights of the sequences among $\mathcal{T}$; the height of each sequence itself is one greater than all sequences that properly extend it. In particular, sequences that have no proper extensions have height 0.

For example, the height of the tree in figure 6 is four. [17]

Why have I introduced this machinery? It allows us to re-define groundedness in a way that captures its *downwards* characterization from p. 25. Something, we may say, is $\gimel$-grounded in gg if it has a $\gimel$-gg-*ww*-priority tree. Thus, something is called grounded in gg if tracing down its generation bottoms out in them. We can show that this definition of groundedness is equivalent to the upwards definition from the previous section.

**Proposition 1** (Aczel 1977, prop. 1.4.5; Yablo 1982, prop. 12)**.** *Let $\gimel$ be some generator and let gg be some things. Then x is $\gimel$-grounded in gg just in case for some well-ordered ww, x has a $\gimel$-gg-ww-priority tree.*

*Proof.* I recast Aczel's proof in the present plural setting. $x$ is $\gimel$-grounded in gg just in case for some well-ordering *ww* and some *w* among *ww*, $x$ is among $I^{w}_{\gimel}(gg)$. The proposition therefore follows directly from the following lemma. □

**Lemma 1.** *Let $\gimel$ be some generator and gg some things. Let w be any one of some well-ordered ww. Then $x \propto I^{w}_{\gimel}(gg)$ just in case x has a $\gimel$-gg-ww-tree of height $\leqslant w$.*

17 Although intuitive, this concept of height is not without subtlety, cf. Hazen [1981].

*Proof.* The claim is proved by an induction on the well-ordered $ww$. For the purpose of this proof, I will write '$v \lll w$' if $v$ precedes $w$ in the ordering $ww$.

Firstly, the induction base. Let $w_0$ be the least of $ww$. In the left-to-right direction, assume that $x$ is among $I^{w_0}_{\daleth}(gg) = gg$. Then by clause 3b of definition 3, $\langle x \rangle$ itself is a priority tree of $x$ of height $w_0$, as required. In the right-to-left direction, let $\mathfrak{T}$ be $x$'s $\daleth$-$gg$-$ww$-priority tree of height $w_0$. Hence it must be $\langle x \rangle$ and by clause 3b again $x \propto gg$, hence $x$ is among $I^{w_0}_{\daleth}(gg)$.

Now for the induction step. Let $w$ be some index among $ww$, and assume that the claim holds for all $v \lll w$. In the left-to-right direction, assume that $x$ is $\daleth$-grounded in $gg$ by $w$-many steps of $\daleth$-generation. Consider those $yy$ from which $x$ is $\daleth$-generated. Each of them is $\daleth$-grounded in $gg$ by less than $w$-many steps. By our induction hypothesis, each $y$ among $yy$ has a $\daleth$-$gg$-$ww$-tree of height less than $w$. Consider all these sequence, their extensions to the left by $x$, and $\langle x \rangle$. By definition 3, all these sequences are a $\daleth$-$gg$-$ww$-tree of $x$ of at most height $w$.

For the right-to-left direction, assume that $x$ has a $\daleth$-$gg$-$ww$-tree $\mathfrak{T}$ of height $v \lll w$ or $w$. Let $\mathfrak{T}^z$ be the sequences $\langle z, \overrightarrow{u_n} \rangle$ for $n \geqslant 0$ and some $z$ such that $\langle x, z, \overrightarrow{u_n} \rangle$ is among $\mathfrak{T}$. We note that $\mathfrak{T}^z$ are a $\daleth$, $gg$-tree of $z$, of height less than $w$. By our induction hypothesis, therefore, $z$ is among $I^{v}_{\daleth}$, $v \lll w$ or $v = w$. Since by assumption, $x$ is $\daleth$-generated from $zz$, we have that $x$ is $\daleth$-grounded in $gg$ by at most $w$-many steps of $\daleth$-generation, as desired. □

Now recall the concept of a grounded object's *rank* (defn. 1). If $x$ has $\daleth$-$gg$-$ww$-rank $w$ then $x$ is not found at any lower stage. By the right-to-left direction lemma 1, it therefore does not have a tree of height $v$ for any $v$ $ww$-earlier than $w$. Consequently, the $\daleth$-$gg$-*rank* of an object with respect to some well-ordering $ww$ is the height of the shortest (least high) $\daleth$-$gg$-$ww$-tree of it.

Now I return to the task of defining mediate priority, or *dependence*. Recall that we say that $x$ is immediately $\daleth$-prior to $y$ if $x$ is among the things from which $y$ is $\daleth$-generated. Intuitively, the corresponding mediate notion of $\daleth$-priority is that of $y$ being involved in step by step $\daleth$-generation of $x$. This, however, makes sense only given some well-ordering, along which $\daleth$-generation is iterated. So we say that $y$ is mediately $\daleth$-prior to $x$ if $y$ is involved in the $\daleth$-generation of $x$ along some well-ordering $ww$. Generation from what? We also need to specify the things $gg$ that we start out from. $\daleth$-dependence will therefore be relative to both a well-ordering and some $gg$.

Given $ww$ and $gg$, the following may appear as a good definition of $\daleth$-dependence. $x$ depends on $y$ if $y$ occurs in some sequence of some $\daleth$-$gg$-$ww$-priority tree of $x$. However, this definition fails to capture

$$\begin{array}{c} \overline{0} = \overline{0} \lor \neg \forall x (x \neq x) \\ | \\ \neg \forall x (x \neq x) \\ | \\ \neg \overline{0} \neq \overline{0} \\ | \\ \overline{0} = \overline{0} \end{array}$$

Figure 7: A Detour **W**-tree

the intuitive notion of $y$ being involved in the generation of $x$, because it blurs significant differences in how things are generated.[18]

Recall the Tarski truth generator **W** of the previous section. It allows us to generate a disjunctive truth $\phi \lor \psi$ from one or the other disjunct. However, when we start with some true literals, and step by step generate all sentences **W**-grounded in them, the disjunct will be generated from one of them first – namely that one which is itself generated earlier than the other disjunct. This fact about groundedness is not captured by the proposed definition. To see this, let $t$ be some term of our language, and consider the **W**-grounded truth $t = t \lor \neg \forall x (x \neq x)$. It is generated from the true literal $t = t$. However, it also has a **W**-priority tree where $\neg \forall x (x \neq x)$ is the single node immediately below it (see figure 7). The proposed definition of $\gimel$-dependence blurs the difference between these two trees. According to it, the disjunction would depend on the universal quantification even though the latter is not involved in the generation of the former from the true literals.

I prefer to work with a definition that tracks such fine details of generation. I would like $\gimel$-$gg$-*ww*-dependence to be sensitive to the order in which the things $\gimel$-grounded in $gg$ are generated from them. This order, however, is readily available, in the $\gimel$-$gg$-*ww-rank* of each thing grounded. Thus, we arrive at the following definition of $\gimel$-$gg$-*ww*-dependence.

**Definition 5.** If $x$ is $\gimel$-generated from $gg$ along some well-ordered *ww*, and its rank is $w$, let us say that $x$ $\gimel$-$gg$-*ww-depends* on $y$ (in symbols: '$y <_{\gimel\text{-}gg\text{-}ww} x$') iff there is a sequence $\langle x, \ldots, y \rangle$ in some $\gimel$-$gg$-priority tree of $x$ of height $w$.

Why is dependence of some thing on another not merely a matter between them, but also involves some $gg$ of which neither may be one? The reason is this: $x$ depends on $y$ if the generation of $x$ goes through $y$. Generation, however, starts somewhere. It makes sense to speak of $x$'s generation only as generation from some $gg$ (possibly

18 I thank Jon Litland for urging me to render this clear.

nothing). Moreover, what matters is *iterated* generation from $gg$, that is generation along some well-ordering. Therefore, it makes sense to speak of $x$'s dependence on $y$ only as in terms of generation from some $gg$ along some well-ordered $ww$. Accordingly definition 5 renders dependence relative to some $gg$ and some $ww$.

Intuitively, $x$ $\daleth$-$gg$-$ww$ depends on $y$ if $y$ occurs in some $\daleth$-$gg$-$ww$ priority tree as $ww$-high as how $ww$-long it takes to $\daleth$-generate $x$ from $gg$. For example, the sentence $\neg p \vee q$ $\mathbf{P}$,$p, q$-$1, 2$-depends on $p$. By lemma 1, the $\daleth$-$gg$-*rank* of an object with respect to some well-ordered $ww$ is the height of its smallest $\daleth$-$gg$-$ww$-tree. Hence, $y <_{\daleth\text{-}gg\text{-}ww} x$ iff some $\daleth$-$gg$-$ww$-tree of minimal height contains a sequence $\langle x, \ldots, y\rangle$.

Thus, a generator $\daleth$ induces a relation of strict grounding, or full priority, and a relation of dependence, or partial priority. As I will argue in later chapters, these *priority relations* bear on the philosophical significance of groundedness. For this, the following proposition is of central importance.

**Proposition 2.** *For all generators $\daleth$, any things $gg$ and any well-ordered things $ww$, the relation of mediate partial $\daleth$-priority $<_{\daleth\text{-}gg\text{-}ww}$ is a well-ordering on the things $\daleth$-grounded in $gg$.*

*Proof.* Let $\daleth$ be a generator, $gg$ be some things and $ww$ some well-ordered things. Everything $\daleth$-$gg$-$ww$-grounded has a unique rank among $ww$. Hence, for all $\daleth$-$gg$-$ww$-grounded things, their ranks are well-ordered. Thus, lemma 2 below ensures an isomorphism between the well-ordering of ranks and the relation of dependence on the grounded things. Hence, $\daleth$-$gg$-$ww$-dependence is a well-ordering on them. □

**Lemma 2.** *For all generators $\daleth$, any things $gg$ and any well-ordered things $ww$, if $x <_{\daleth\text{-}gg\text{-}ww} y$ then the $\daleth$-$gg$-$ww$-rank of $x$ is strictly smaller than that of $y$.*

*Proof.* Assume that in some $\daleth$-$gg$-$ww$-tree $\mathcal{T}$ of $y$ there is a sequence $\langle y, \ldots, x\rangle$, and the height of $\mathcal{T}$ is the $\daleth$-$gg$-$ww$-rank of $x$. Now assume, for contradiction, that the rank of $x$ is not strictly smaller than that of $y$. By lemma 1, $x$ has no $\daleth$-$gg$-$ww$-tree of height smaller than the $\daleth$-$gg$-$ww$-rank of $y$. Now, I construct from $\mathcal{T}$ a $\daleth$-$gg$-$ww$-tree of height smaller than the $\daleth$-$gg$-$ww$-rank of $y$, showing that $x$'s rank must in fact be strictly smaller than $y$'s.

So take all sequences among $\mathcal{T}$ that contain $x$, and from each of them, chop off its initial segment up to the first occurrence of $x$. Call the resulting collection of sequences $\mathcal{T}'$. I claim that the result is a $\daleth$-$gg$-$ww$-tree of $x$ strictly smaller than the $\daleth$-$gg$-$ww$-rank of $y$.

Recall definition 3 and note firstly that among $\mathcal{T}$ there are sequences whose last item is $x$. $\mathcal{T}'$, therefore, contains $\langle x\rangle$. It is the only sequence of length one, because all sequences in $\mathcal{T}'$ are ensured to begin with $x$. Thus, $x$ is the root of $\mathcal{T}'$. Secondly, because the sequences in $\mathcal{T}$ satisfy clauses 2 and 3 of definition 3, so do those in $\mathcal{T}'$.

Finally, the fact that $\mathcal{T}$ satisfies clause 3 equally carries over to its subtree $\mathcal{T}'$, since we only chopped off initial segments of sequences among $\mathcal{T}$. I conclude that $\mathcal{T}'$ is a $\beth$-$gg$-$ww$-tree of $x$. However, it is a proper subtree of $\mathcal{T}$, hence shorter, so $x$ after all has a $\beth$-$gg$-$ww$-tree of height strictly smaller than the $\beth$-$gg$-$ww$-rank of $y$, as desired. $\square$

So far, the concepts of groundedness and dependence have always been relativized to some given well-ordered $ww$. However, this will prove inconvenient in the long run. Fortunately, though, it is also often unnecessary. Most well-orderings considered are of set sized order type. In these cases, we can always resort to the ordinal numbers less than it, and let them witness the assumed well-ordered $ww$. Therefore, in practice I will often be able to suppress mention of them. If I do so, the relevant well-ordered things are a long enough initial segment of the von Neumann ordinals.[19]

## 2.5 VARIETIES OF GENERATION

In this section, I present general concepts and simple results about them, thus building up a set of tools for my later, more specific investigations. The impatient reader may skip ahead to applications (§§ 2.6ff.)

Given a generator $\beth$, it may be the case that $\beth$ allows for the construction of one and the same object $x$ from distinct pluralities $yy$ and $zz$. Below I will give reason to focus on cases in which this is ruled out. I will consider generators $\beth$ that are *left-unique*, in the following sense.[20]

**Definition 6** (Left- and right-uniqueness)**.** Let us call a generator $\beth$ *left-unique*, iff for every $x$ and all pluralities $yy, zz$

If $yy\beth x$ and $zz\beth x$ then $yy$ are $zz$.

I call a generator $\beth$ *right-unique*, iff for all pluralities $xx$ and every $y, z$,

If $xx\beth y$ and $xx\beth z$ then $y$ is $z$.

The generator **P** from example 2.2 is left-unique, but not right-unique. $(p \vee q)$ is generated from precisely $p$ and $q$, from which, however, we may also generate $(p \wedge q)$.

One reason to be interested in *left-unique* generators is that for left-unique $\beth$ and some $gg$, every object $x$ has a *unique* (up to isomorphism) $\beth$, $gg$-tree.[21]

**Lemma 3.** *Let $\beth$ be a* left-unique *generator, let $x$ be some object and let $ww$ be well-ordered. For every $gg$, $x$ has exactly one $\beth$-$gg$-$ww$-tree (up to isormophism).*

19 This convention of course does not apply to the groundedness of the ordinal numbers themselves, and the groundedness of pure sets, because in these cases, there is no large enough ordinal.

20 Occasionally, such generators are called 'deterministic' [Aczel, 1977, p. 744].

21 This basic observation is implicit in Forster's 2008 discussion.

*Proof.* Let $\mathcal{T}$ and $\mathcal{T}'$ be $\gimel$-$gg$-$ww$-priority trees of $x$. I show that $\mathcal{T} = \mathcal{T}'$ by induction on the height of $\mathcal{T}$. Since it is finite, I can represent the arbitrary $ww$ by the natural numbers.

If $\mathcal{T}$ is the single sequence $\langle x \rangle$, then by clause (4b) of definition 3, $x$ must be among $gg$. Hence, for $\mathcal{T}'$ to be a $\gimel, gg$-tree of $x$ it must firstly have $x$ as its root, and its longest sequences must end in $gg$. $\mathcal{T}'$ must therefore be the single sequence $\langle x \rangle$, and thus be identical to $\mathcal{T}$.

Now let $\mathcal{T}$ be of height $n+1$, and let $\mathcal{T} \upharpoonright n$ be $\mathcal{T}$'s largest subtree of height $n$ (for intuition, $\mathcal{T} \upharpoonright n$ is $\mathcal{T}$ without its leaves). Since $\mathcal{T}$ is a $\gimel, gg$-tree, we know that for every $\mathcal{T} \upharpoonright n$-sequence of length $n$, its rightmost item (a '$\mathcal{T} \upharpoonright n$-leaf') is $\gimel$-generated from some $\mathcal{T}$-leafs. Similarly, all $\mathcal{T}' \upharpoonright n$ are $\gimel$-generated from $\mathcal{T}'$-leafs.

Now assume for contradiction that $\mathcal{T} \neq \mathcal{T}'$. By our induction hypothesis we know that $\mathcal{T} \upharpoonright n = \mathcal{T}' \upharpoonright n$. So, $\mathcal{T}$ and $\mathcal{T}'$ must differ on their leaves. Hence, some leaf of the subtree $\mathcal{T} \upharpoonright n = \mathcal{T}' \upharpoonright n$ must be generated from some $zz$ distinct from those $gg$ that it is generated from in $\mathcal{T}'$. This, however, contradicts our assumption that $\gimel$ is *left-unique*. $\square$

Lemma 3 ensures that the corresponding relation $\leqslant_\gimel$ of being grounded in (definition 1) is non-monotone in the sense that if $xx \leqslant_\gimel y$ then there is no $zz \sqsupseteq xx$ such that $zz \leqslant_\gimel y$.

Let me draw a few other simple distinctions among generators, and make some basic observation that to the best of my knowledge have not yet been made explicit.

**Definition 7** (Closure)**.** Let us call a generator $\gimel$ *left-closed* iff for every $x$ and all pluralities $yy, zz$

> If $yy \gimel x$ and $zz \gimel x$ then $yy, zz \gimel x$.

For example, **P** is left-closed.

**Definition 8** (Cover)**.** Let us call a generator $\gimel$ *covered* iff for every $x$ and all pluralities $yy, zz$

> If $yy \gimel x$ and $zz \gimel x$ then for some $uu$ such that $xx, zz \sqsubseteq uu$, $uu \gimel x$.

Note that uniqueness implies left-closure, which in turn implies cover.

Is left-closure preserved when generators are combined? Not in general. To see this, let $\gimel$ and $\daleth$ be two left-closed generators such that $xx \gimel y$, but not $zz \gimel y$, and $zz \daleth y$, but not $xx \daleth y$. For $xx, zz \gimel$-$\daleth y$ we need that either $xx, zz \gimel y$ or $xx, zz \daleth y$, neither of which, however, is ensured by the left-closure of the generators $\gimel$ and $\daleth$.

Such considerations motivate to look more closely at combinations of generators.

**Definition 9** (Interference)**.** We say that generators $\gimel$ and $\daleth$ *interfere* iff there is an $x$ such that for some $yy, zz$,

$yy \gimel x$ and $zz \daleth x$

Now, we can observe that the combination of two left-closed generators is left-closed if they do not interfere.

An analogous observation can be made about left-uniqueness.

**Definition 10** (Divergence)**.** We say that generators $\gimel$ and $\daleth$ *diverge* iff for some $xx$ there are distinct $y$ and $z$ such that

$xx \gimel y$ and $xx \daleth z$

The combination of two left-unique generators is itself left-unique if they do not diverge.

In the remainder of this chapter I present two prominent and philosophically significant cases of groundedness.

## 2.6 CANTORIAN NUMBERS

In his *Grundlagen*, Cantor presents the ordinal numbers, his *extended number sequence*, as those obtained by two *principles of generation* [1932, pp. 195f]. Firstly, given a number we generate its successor. Secondly, [Ewald, 1996, pp. 907f]

> [. . . ] if any definite succession of defined integers is put forward of which no greatest exists, a new number is created [. . . ], which is thought of as the *limit* of those numbers; that is, it is defined as the next number greater than all of them.

Two comments are in order. Firstly, both principles appear to presuppose a way in which the numbers are ordered. For example, to apply the first principle, it seems, we need to know already which number succeeds which. I follow Jané [2010, p. 197] and understand Cantor as taking this ordering of numbers to be just the order in which they are generated. Cantor's principles do not presuppose the ordering of the numbers, but provide it themselves. It is not as if the first principle lets us, so to speak, pick from the given order of numbers the next one. Instead, given a number $x$ it lets us generate a new number $y$, and in virtue of having been generated this way, $y$ is the successor of $x$.

Analogously, it is strictly speaking not the case that the second principle allows us to generate, for any definite sequence of numbers, their least upper bound – rather, it allows us to generate a number, and doing so to extend the ordering by a least upper bound. Accordingly, I understand the term 'succession' in the above quote, and 'sequence' as in 'extended number sequence', as referring to precisely this Cantorian order of generation.

Secondly, how exactly to spell out Cantor's notion of definite collections is subject to scholarly debate. Without attempting to do justice to its breadth, let me remark that 'collection' may be understood on a par with 'plurality', as convenient shorthand for essentially plural terminology, such that it principle it can in all contexts be paraphrased in purely plural terms (recall my fn. 5). What does it mean for a plurality to be definite? In the present context, an answer to this question must be informed by the Burali-Forti paradox. Notoriously, Cantor's second principle is incompatible with the assumption of a definite collection of all ordinals. Assume they formed a definite collection. Then, by the second principle there is an ordinal greater than all ordinals, contradiction.

The Burali-Forti paradox thus poses constraints on how to spell out Cantor's notion of definiteness. Broadly speaking, there are two routes. On the one hand, we may accept that any plurality is definite, but reject the view that every condition defines a plurality (see chapter 4 below, especially p. 64). On the other hand, we may hold that for every condition, there are exactly those things which satisfy it, but that not every plurality is definite, in particular that the plurality of all ordinals fails to be such.

Each route has its advantages, but also comes at significant costs, too. I do not need to take a stance on this issue, but may defer to the relevant literature for a safe explication of definiteness [Dummett, 1978; Shapiro and Wright, 2006]. Taking my first and second comment together, I understand Cantor's second principle as follows: Given any *definite* plurality of numbers, a new number is generated that is not one of but greater than all of them.

Now, I capture the Cantorian generation of ordinal numbers by generators in the sense of section 2. This will allow me to view the ordinal numbers as grounded. Naturally, the first principle is captured by the generator which given a number, outputs its successor; the second principle is captured by one which given some numbers, outputs their least upper bound.

**Definition 11** (Cantor's Number Generators)**.** Given any definite, possibly empty, plurality of numbers $\alpha\alpha$,

1. $\beta$ is **C1**-generated from them iff $\alpha\alpha$ are exactly one ordinal $\alpha$, and $\beta$ is $\alpha$'s successor.

2. $\beta$ is **C2**-generated from them iff $\beta$ is their least upper bound

In particular, the least number is **C2**-generated from nothing. I repeat, the restriction to definite pluralities in this definition is subject to one's chosen response to the Burali-Forti paradox. The restriction comes out as vacuous if every plurality is definite, in which case we must not assume every condition to define a plurality. Alternative approaches, however, require a non-vacuous concept of definiteness, which may then be, so to speak, plugged into the above definition.

At this point, it proves advantageous that I formulated groundedness without reference to ordinals, but in terms of arbitrary well-orderings understood plurally (definition 1 on p. 26). An ordinal itself is grounded, by combination of **C1** and **C2**, and along an ordering whose type it is.

**Proposition 3.** *Assume that some definite things ww are well-ordered, and that the ordinal $\alpha$ is their order-type. Then $\alpha$ is **C1**-**C2**-grounded in nothing.*

*Proof.* By induction on $\alpha$. If it is 0, it is **C2**-generated from nothing. If it is a successor ordinal, it is **C1**-generated from its predecessor, which is **C1**-**C2**-grounded in nothing by the induction hypothesis. A limit ordinal is **C2**-generated from the ordinals smaller than it, hence **C1**-**C2**-grounded in nothing, too. □

In the literature, $\gimel$-groundedness is usually defined by the least fixed point of the operator $\Gamma_{\gimel}$ (p. 28). I explained how this concept is recovered within my framework. However, I deliberately chose a different definition: groundedness is being *at some stage* of iterated $\gimel$-generation. The present case of Cantorian number groundedness justifies this decision. It allows me to present the Cantorian notion of the ordinal numbers as **C1**-**C2**-generated from nothing, without committing myself to a least plurality closed under ordinal generation, that is, to a plurality of all ordinals.

## 2.7 THE WELL-FOUNDED SETS

I have developed Forster's [2008] generalized iterative conception into a theory of groundedness. His starting point and primary example is the *cumulative hierarchy of sets*. Therefore, it is apposite to explain how my general theory applies to the well-founded sets.

The cumulative hierarchy of sets plays an important role in the philosophy of set theory. According to a widely held view, all and only the sets there are, are those in the cumulative hierarchy. This *iterative conception* of sets has a venerable history.[22] Possibly the first and arguably the most influential formulation is due to Gödel [1947, p. 180]. He characterizes the iterative conception of set as that view

> [...] according to which a set is anything obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of" [...].

Gödel's set-of operation is captured well by a generator that turns some things into their set. For example, the singleton $\{\varnothing\}$ would thus be generated from its single element, the empty set. However, some care is needed when defining this generator. Not from any collection of things we can generate their set, witness those sets which are not elements of themselves. Just as in the previous section the Burali-Forti paradox required a restriction of the Cantorian generators to

22 For a recent, opinionated overview see Ferreiros [1999, p. 441 - 456].

definite pluralities of ordinals, Russell's paradox now requires some restriction of the Gödelian set generator to those collections of things whose sets can be formed consistently.

As before, however, I do not need to take a stance on how precisely to identify these set forming pluralities. Maybe, every plurality forms a set, and there simply are no things which satisfy Russell's condition of not being an element of itself. Or, there is such a plurality, but it does not form a set. Using 'definite' in the same parametrical manner as in the previous section, my treatment of the paradox is put as follows. Maybe every plurality is definite, or maybe Russell's paradox requires set formation be restricted to definite pluralities. Either way, it is safe to define the Gödelian set generator in the following manner.

**Definition 12** (The set generator)**.** Let $xx$ be a definite, possibly empty, plurality of things. $y$ is **S**-generated from $xx$ if $y$ is the set of $xx$, that is, $xx$ are the elements of $y$.

In terms of **S**, we can paraphrase Gödel's statement of the iterative conception as follows: A set is anything **S**-grounded in the integers or some other urelemente. Nowadays, it is more common to focus on the pure sets. This iterative conception of pure sets is equally well expressed in my framework: A pure set is anything **S**-grounded in nothing (also recall figure 1 on p. 16). From now on, I focus on the pure sets and by '**S**-groundedness' always mean **S**-groundedness in nothing.

The notion of **S**-groundedness is no stranger to set theory. Quite the contrary, it is long and well known, if only under a different label. A pure set is **S**-grounded if and only if it is *well-founded*. To see this, recall definition 2 on p. 28 and note that **S** gives rise to an operator $\Gamma_{\mathbf{S}}$ which takes some things $xx$ and outputs all sets formed from their definite sub-pluralities.

$$y \propto \Gamma_{\mathbf{S}}(xx) :\Leftrightarrow \exists zz \sqsubseteq xx(y = \{zz\}) \tag{9}$$

Thus, $\Gamma_{\mathbf{S}}$ is a plural power-set operation. A set $x$ is **S**-grounded in nothing if starting from nothing, and iterating the operator $\Gamma_{\mathbf{S}}$, we eventually arrive at $x$.

Since some things form a set, so do its subsets (assuming both pluralities are definite). Consequently, if $xx$ form a set, $\Gamma_{\mathbf{S}}(xx)$ form its power-set.[23] Now, **S** allows us to generate the empty set from nothing. Hence, $\Gamma_{\mathbf{S}}$ applied to nothing gives some things that form a set. So, each stage of iterating $\Gamma_{\mathbf{S}}$ form a set, namely some initial segment $V_\alpha$ of the cumulative hierarchy. Hence, $x$ is **S**-grounded in nothing if and only there is some $\alpha$ such that $x \in V_\alpha$.

23 Note that here, my choice of a plural meta theory pays off. Forster states his generalized iterative conception in standard set theoretic term, which blurs the important difference between some things and the set which is **S**-generated from them.

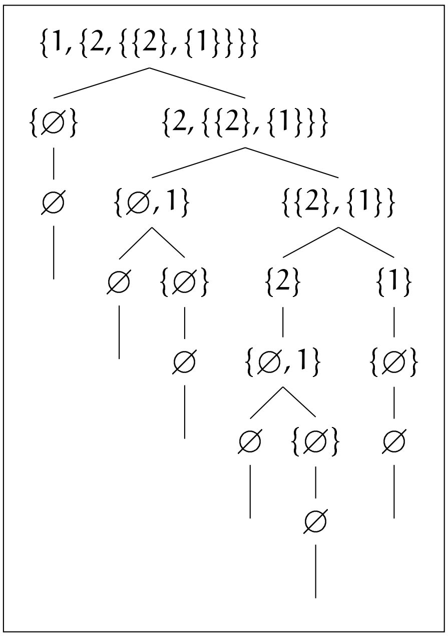


Figure 8: An **S**-priority tree (numerals abbreviate von Neumann ordinals)

**Proposition 4.** *For all* $x$*,* $x$ *is* **S***-grounded in nothing if and only if* $x$ *is a pure, well-founded set.*

Thus, to be **S**-grounded is to be a set of the cumulative hierarchy. This should not surprise. After all, the formal concept of groundedness from the previous sections builds on Forster's generalization of the cumulative hierarchy [2008].

By the general proposition 1, if $x$ is **S**-grounded then it has an **S**-priority tree. This **S**-priority tree of $x$ gives its elements, their elements, and so on. Figure 8 shows the **S**-priority tree of $\{1, \{2, \{\{2\}, \{1\}\}\}\}$ (using numerals to denote the von Neumann ordinals). Infinite sets, of course, will have trees with infinitely many nodes.

This observation allows us to connect the concepts of chapter 2 with standard terminology from set theory. Note that **S** is left-unique (definition 6) – if $yy\mathbf{S}x$ and $zz\mathbf{S}x$ then $yy$ are $zz$. By lemma 3, therefore, every set has a unique **S**-priority tree. Now, this tree of $x$ is isomorphic to the *transitive closure* of $x$, ordered by set elementhood. By proposition 4 we therefore know that **S**-dependence (def. 5) is well-founded.

Thus, my general concept of groundedness from the previous section provides a neat formalization of the iterative conception of sets. In order to show that it is capable of more, I will in the next chapter apply it to Kripke's concept of semantic groundedness. Later in the present thesis, however, I will return to **S**-groundedness when developing my account of the philosophical significance of groundedness (chapter 6).

## 2.8 CONCLUSION

In this first chapter of my thesis, I introduced my formal concept of groundedness, proved central propositions about groundedness, and applied my general theory to three prominent examples. The primitive concept of my theory is that of a generator. In terms of it, I gave two formal definition of groundedness, each intended to capture a distinct intuition. Then, improving on an early result by Yablo [1982], I showed them to be equivalent. Along the way, I presented the case of groundedness based on the Tarskian truth generator **W**.

In section 2.5 I continued with certain, still fully general, categories of generation and groundedness. In the remaining two sections, I then applied my theory to firstly, the Cantorian view of the ordinals as being generated according to two principles, and secondly to the well-founded sets. Both cases proved to be simple and illuminating instances of groundedness. In the next chapter, I will turn to a more complicated but philosophically very interesting case of groundedness: Kripke's semantic groundedness.

# 3

# GROUNDED TRUTH

## 3.1 INTRODUCTION

Consider the language $\mathcal{L}_{ta}$ of first-order arithmetic extended by a predicate symbol 'T'. For simplicity, I assume '$\neg$', '$\vee$', '$\forall$', '$S$' and '$\overline{0}$' to be all the primitives in terms of which the other connectives, quantifiers and arithmetical symbols are defined. I fix, once and for all, some reasonable method of associating every sentence $\phi$ with its *Gödel code* $\ulcorner\phi\urcorner$.[1] I will use capital Roman letters from the end of the alphabet ('X', 'Y' etc.) as ranging over sets of $\mathcal{L}_{ta}$-sentences, or sets of sentence codes, depending on the context.

In his seminal 1975 article, Kripke showed how to expand the standard model of arithmetic $\mathfrak{N}$ by interpretations X of 'T' of particular philosophical interest. The core of his construction is the *Kripke jump* from truth in a model $\mathfrak{M}(X)$ to a new interpretation Y of 'T', and thus to a new model $\mathfrak{M}(Y)$.[2] X is a Kripke truth predicate if it is a fixed point of such a jump.

My interest is in a certain kind of Kripke truth predicates, those known as predicates of *grounded* truth. The notion of grounded truth is due to Hans Herzberger 1970, but in his 1975 paper, Kripke provides it with new and original content. He does so by telling a story how an idealized speaker comes to know the concept of truth [Kripke, 1975, pp. 701ff]. This story has been retold many times since [Visser, 1983; McGee, 1991; Maudlin, 2004]. Nonetheless, I will present it once more, because it renders vivid the close kinship of Kripke's semantic groundedness with the groundedness of numbers and sets; and this aspect of the famous story has not been sufficiently recognized yet.

## 3.2 LEARNING THE TRUTH

Consider Alice. She speaks a peculiar fragment of English: English except for the word 'true'. Further, let us, as usual in philosophy, idealize and assume Alice to have unlimited cognitive capacities and to know for every proposition expressible in this language, whether or not it is true.

Now we present to Alice English sentences with 'true'. She does not understand them, because she does not know the meaning of this word. So we tell her to call a proposition 'true' just in case that she can assert it, and 'not true' whenever she is entitled to deny it. She already knows that, for instance, snow is white. Recognizing that she can assert this proposition, Alice applies the rule she has just been given and infers that she can also say that 'snow is white' is true. Similarly, she proceeds with everything else that she already

1 To be precise, '$\ulcorner\phi\urcorner$' will denote the Gödel number of $\phi$ or its Gödel *numeral*, depending on the context.

2 As usual, '$\mathfrak{M}(X)$' denotes the *expansion* of the model $\mathfrak{M}$ by a relation X on the domain of $\mathfrak{M}$. This differs from the *extension* of a model in that the domain does not increase.

knows. Due to her perfect knowledge of non-semantic facts and her remarkable cognitive capacities, this means that for every proposition *expressible in English minus 'true'*, she has now come to understand every sentence in which 'true' applies to a term for this proposition.

Now, Alice again applies the rules that we have given her, this time to these newly understood sentences. Step by step, she therefore learns to apply 'true' to more and more sentences, also to those which contain 'true' themselves. Since Alice is an idealized subject, it is appropriate to assume that she iterates this step along all natural numbers, and reaches a first limit stage. Here, she takes stock of what she has learnt, and understands that she can apply her new word 'true' to every true propositions that she can express so far. Then, Alice continues to learn more sentences, step by step.

Let us focus on how the extension of 'true' increases during this process (see figure 9). First, 'true' applies to nothing at all. Then, Alice understands that it applies to every true sentences without 'true'. At the third stage of her learning process, she comes to know that 'true' also applies to every true sentence in which 'true' is applied to a sentence not containing 'true' itself. At limit stages, the extension of 'true' is the union of all previous stages.

Kripke suggests that 1975, p. 701

> [...] the "grounded" sentences can be characterized as those which eventually get a truth value in this process'

In other words, sentence is grounded if and only if at some stage, it enters the extension of 'true'. Kripke then gives a general method of constructing models that capture this notion of semantic groundedness. Kripke's construction is well known, and there are several excellent presentations of it [McGee, 1991; Horsten, 2011]. Nonetheless, certain aspects of it have not received sufficient attention. They are interesting in their own right, but will also become relevant to my treatment of semantic groundedness in later chapters. In particular, from the available presentations it is not obvious how Kripke's concept of semantic groundedness instantiates the previous chapter's general concept of groundedness. In the remainder of this chapter, I will therefore recast Kripke's construction and attempt to work out these aspects.

## 3.3 kripke's construction

Starting out from some base theory in the language $\mathcal{L}_a$, usually the set of truths in the standard model of arithmetic $\mathfrak{N}$, the Kripke jump is iterated and more and more sentences containing 'T' enter its interpretation. Kripke calls a sentence "grounded" if it or its negation enters the interpretation at some stage of this construction. The least fixed point extending the base theory collects all and only the grounded sentences.

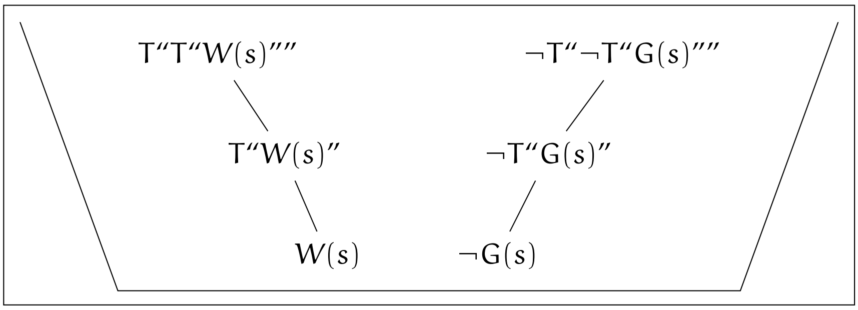


Figure 9: How Alice learns Truth

For this set to be consistent, however, “truth in $\mathfrak{N}(X)$” must not mean classical satisfaction. Assume it did. For any sentence $\phi$ containing ‘T’, the jump of our base theory would contain $\neg T\ulcorner\phi\urcorner$, even if later on, $\phi$ enters the truth predicate and $T\ulcorner\phi\urcorner$ comes out as true. For example, while $0 = 0$ is a sentence of arithmetic, $T\ulcorner 0 = 0\urcorner$ is not. Therefore, $\neg T\ulcorner T\ulcorner 0 = 0\urcorner\urcorner$ would be found at the first stage. Likewise, however, this first stage would contain the sentence $T\ulcorner 0 = 0\urcorner$, since $0 = 0$ is among our base theory. Jumping ahead just once, we would obtain $T\ulcorner T\ulcorner 0 = 0\urcorner\urcorner$, the very sentence whose negation we have just found grounded. Consequently, a Kripke jump based on classical satisfaction creates inconsistent truth predicates.

The reason is that classical satisfaction lets $\neg T\ulcorner\phi\urcorner$ come out true whenever $\phi$ is not in the interpretation of ‘T’. Hence, Kripke’s construction must not be carried out on the basis of classical satisfaction, in particular not on the basis of the classical treatment of *negation*. Speaking loosely, an evaluation scheme is needed that renders a negated sentence true not in the absence of information, but if the available information suffices for it. More precisely, an evaluation scheme $m$ is needed such that $\mathfrak{M}(X) \models_m \neg\phi$ only if $\mathfrak{M}(Y) \models_m \neg\phi$ holds for every interpretation Y of ‘T’ extending X. Using a technical term, satisfaction must be *monotone* [Blamey, 2002]. The crucial feature of a monotone evaluation schema $m$ is that the fact that some sentence code $\ulcorner\phi\urcorner$ is *not* in the extension of ‘T’ no longer suffices for the negation $\neg T\ulcorner\phi\urcorner$ to come out as true. We no longer have that $\neg T\ulcorner\phi\urcorner$ is true in $\mathfrak{N}(X)$ if $\phi$ is not among the sentences X.

Various monotone schemes $m$ have been used for Kripke’s construction. Thus, we have a Kripkean truth predicate based on *Strong* and *Weak* Kleene logic, and constructions that use supervaluational schemes. Note, however, that the need for monotone satisfaction does not imply that our theory of grounded theory cannot be classical. All we have found is that classical logic must not be used for the Kripke jump, if our goal is a consistent truth predicate. The Kripke jump must be formulated using non-classical logic. But, what we do with our truth predicate thus obtained is a different matter. In particular, we may well reason classically with it. Technically, this means we can take a Kripke truth predicate X and work within a classical model $\mathfrak{M}(X)$. This approach goes back to Kripke [1975, p. 715] and has been discussed as "closing off" the non-classical model. I will consider its advantages and disadvantages in the next chapter (p. 70).

Given a monotonic evaluation scheme $m$, the Kripke jump is standardly formalized by an operator $\mathcal{J}_m$ on sets X of (codes of) $\mathcal{L}_{ta}$-sentences.

$$\ulcorner\phi\urcorner \in \mathcal{J}_m(X) \text{ iff } \mathfrak{N}(X) \vDash_m \phi \tag{10}$$

For example, $\mathcal{J}_{sk}$ is the standard Kripke jump based on the *Strong Kleene* scheme $\vDash_{sk}$ (see definition 14 below).

Kripke called a sentence *grounded* if its code is found in the least fixed point of $\mathcal{J}_m$. My goal is to show that this particular concept of semantic groundedness is a special case of the general concept from section 2. For this, I need to provide a *generator* on the $\mathcal{L}_{ta}$-sentences. They form a set *Sent*, which allows me to proceed in the usual set-theoretic setting. In particular, I can represent a generator $\beth$ by a set of pairs $\langle X, \phi\rangle$, where $X \subseteq$ *Sent* and $\phi \in$*Sent*.

Like the generalized power-set operation of section 2.7, the operator $\mathcal{J}_m$ is an example for operators $\Gamma_\beth$ from chapter 2 (p. 28). However, what generator $\beth$ does it correspond to? It is a generator that allows us to infer $T\ulcorner\phi\urcorner$ from a set of sentences X if $\mathfrak{N}(X) \vDash_m \phi$. Given X, we generate all sentences $T\ulcorner\phi\urcorner$ such that $\phi$ is in the Kripke jump of X. Accordingly, I will speak of the *jump* generators and refer to them by **JM**. For example, the Strong Kleene schema $\vDash_{sk}$ gives rise to the jump generator **JSK**. In general, **JM** is given by the following rule.

$$\frac{X}{T\ulcorner\phi\urcorner} \text{ if } \mathfrak{N}(X) \vDash_m \phi \tag{11}$$

In the previous section, I mentioned Yablo's early work on groundedness. Now, I note that a generator **JM** corresponds to Yablo's notion of *jump-entailment*, or *sufficiency*, for a Kripke jump $\mathcal{J}_m$ [Yablo, 1982, p. 121]. Yablo in turn ascribes this concept to Herzberger [Yablo, 1982, fn. 7].

Of course, a generator **JM** allows us also to draw priority trees, as in section 3, and thus gives rise to a corresponding notion of *dependence*. Figure 10 provides one example, based on the Strong Kleene jump generator **JSK**.[3] Note that its root, a negated disjunction of the form $\neg(T\ulcorner\phi\urcorner \vee T\ulcorner\psi\urcorner)$ depends not on what is negated, nor on either disjunct, but directly on the sentences of which truth is predicated.

Presented in this manner, Kripke's concept of semantic groundedness appears to be as simple an instance of the general concept as is the cumulative hierarchy of sets (§2.7 in the previous chapter). A sentence is true in Kripke's least fixed point models if and only if it is **JM**-grounded in nothing.

3 As usual, I abbreviate by $\dot{x}$ a PA-representation of the function that maps a number $n$ to its numeral $\overline{n}$.

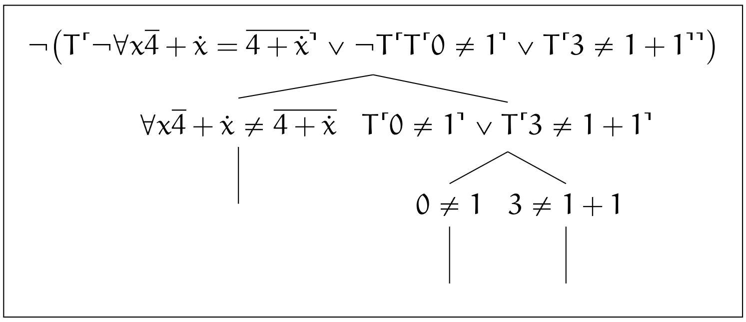


Figure 10: Low resolution Kripke Groundedness: A **JSK** Priority Tree

I will refer to this as the *low-resolution* characterization of semantic groundedness. My reason for this label is the following. If we look more closely at Kripke's fixed point construction, we find a richer structure of interacting groundedness than the above, received characterization suggests. Given some $xx$, their set is obtained directly. Given some sentences $X$, however, its Kripke jump $\mathcal{J}_m(X)$ is better viewed as being obtained in two steps (see figure 11).

Firstly, we ascribe truth to all and only the sentence in $X$, thus moving from complex sentences $\phi \in X$ to atomic sentences $T\ulcorner\phi\urcorner$, and infer $\neg T\ulcorner\phi\urcorner$ if $\neg\phi \in X$. Secondly, this collection of literals is closed under logic.[4] More precisely, the set of literals $\{T\ulcorner\phi\urcorner : \phi \in X\} \cup \{\neg T\ulcorner\psi\urcorner : \neg\psi \in X\}$ is closed under the consequence relation $\models_m$ which corresponds to the monotone evaluation scheme $m$. Doing so, we obtain precisely the sentences $m$-true in the model $\mathfrak{N}(X)$, in other words the Kripke jump $\mathcal{J}_m(X)$. In sum, taking the Kripke jump $\mathcal{J}_m$ of a given set $X$ involves two steps: firstly, we ascribe truth, secondly, we close under the monotone logic $m$. This fact is missed if we understand Kripke's concept of semantic groundedness in terms of generators **JM**, that allow us to move from the sentences $X$ directly to the complete theory of $\mathfrak{N}(X)$. Therefore, generators **JM** provide a merely low-resolution characterization of Kripkean semantic groundedness.

Fortunately, a finer *high-resolution* understanding of it is available. In the next section, I will outline a general method of replacing a single generator **JM** by two generators **T** and **M** that capture the two distinct steps behind the Kripke jump.

## 3.4 SEPARATING TRUTH FROM LOGIC

The first step, moving from the set $X$ to the set of literals $\{T\ulcorner\phi\urcorner : \phi \in X\} \cup \{\neg T\ulcorner\psi\urcorner : \neg\psi \in X\}$, corresponds to the generation of sentences $T\ulcorner\phi\urcorner$ from $\phi$ and $\neg T\ulcorner\phi\urcorner$ from $\neg\phi$. This *truth generator* **T** is common to each variant of Kripke's construction, whichever monotone evalu-

4 As usual, I call a sentence $\phi$ a 'literal' if it is atomic or the negation of an atomic sentence.

ation scheme $\mathfrak{m}$ we choose. In this sense, **T** is the core of Kripke's construction.

**Definition 13** (Truth Generator)**.** Let the generator **T** be given by the following rules, for $\mathcal{L}_{ta}$-sentences $\phi$.

$$\text{T-Intro}\ \frac{\phi}{T\ulcorner\phi\urcorner} \qquad \frac{\neg\phi}{\neg T\ulcorner\phi\urcorner}\ \neg\text{T-Intro}$$

Note that **T** allows us to generate two distinct sentences $T\ulcorner\neg\phi\urcorner$ and $\neg T\ulcorner\phi\urcorner$ from any negation $\neg\phi$. Further, by itself, **T** allows us to generate more and more statements of the form "it is true that .." and "it is not true that . . . ", from some given set of sentences, say the truths of arithmetic. However, we will not arrive at any conjunction, disjunction or quantification of such statements. For this, we need to close the set $\{T\ulcorner\phi\urcorner : \phi \in X\} \cup \{\neg T\ulcorner\psi\urcorner : \neg\psi \in X\}$ under logic. What, however, does it mean to close a set of literals under logic? This depends on our choice of a monotone evaluation scheme $\mathfrak{m}$. In the next section, I will identify *logic* generators **M** such that $\mathfrak{N}(X) \vDash_{\mathfrak{m}} \phi$ iff $\phi$ is **M**-grounded in the literals $\mathfrak{m}$- true in $\mathfrak{N}(X)$.

In combination, **T** and **M** provide us with the following high-resolution characterization of Kripke's concept of grounded truth.[5]

**Proposition 5.** *Let $\mathfrak{m}$ be either the Weak or Strong Kleene evaluation scheme, and **M** be the corresponding logic generator. Then $\ulcorner\phi\urcorner$ is in the least fixed point of Kripke's jump operator $\mathfrak{J}_{\mathfrak{m}}$ if and only if $\phi$ is T-**M**-grounded in the $\mathcal{L}_a$-literals true in $\mathfrak{N}$.*

*If $\mathfrak{m}$ is a supervaluational schema and **M** the corresponding generator, then $\ulcorner\phi\urcorner$ is in the least fixed point of Kripke's jump operator $\mathfrak{J}_{\mathfrak{m}}$ if and only if $\phi$ is T-**M**-grounded in the $\mathcal{L}_a$-sentences true in $\mathfrak{N}$.*

The proposition follows from lemmata given in the next sections. In the next sections, I will prove the lemmata sufficient to establish the proposition (lemmata 5, 8 and 6 below).

My conclusion of the foregoing discussion is that within the general framework of section 2 there are two ways of understanding Kripke's concept of semantic groundedness. On the one hand, there is the standard, *low-resolution* characterization. Given some grounded truths X, more are generated by taking the Kripke jump of X. In particular, $T\ulcorner\phi\urcorner$ is generated not from $\phi$ alone but from *all* sentences that have already entered the interpretation of 'T'. On this characterization of semantic groundedness, it becomes a rather simple instance of the general theory of groundedness from section 2, **JM**-groundedness in nothing (equation 11 above).

On the other hand, there is the *high-resolution* notion based on the combination of a uniform truth generator **T** with one of the logic

5 There is a connection between this proposition and Kit Fine's 2010 presentation of Kripke's theory. I assume that he had arrived at a similar result.

$$\begin{tikzcd} \zeta_0,\zeta_1,\dots,\neg\xi_0,\neg\xi_1,\dots \arrow[d, "\mathbf{T}"] \arrow[dr, "\mathcal{J}_m"] & \\ T\ulcorner\zeta_0\urcorner,\dots,\neg T\ulcorner\xi_0\urcorner,\dots \arrow[r, "\mathbf{M}"] & \psi_0,\psi_1,\dots,\neg\chi_0,\neg\chi_1,\dots \end{tikzcd}$$

Figure 11: Splitting Kripke's jump into truth-generation **T** and closure under some monotone logic generator **M** .

generators **M**. On this reading, Kripke's move from a set $X$ to its Kripke jump falls into two steps. Firstly, the sentences $X$ are ascribed truth, and if $\neg\xi \in X$ then $\xi$ is inferred not to be true. We obtain a set of literals $T\ulcorner\xi\urcorner$, $\neg T\ulcorner\xi\urcorner$. It is only in a second step that this set of sentences is closed under the logic determined by our chosen evaluation scheme $m$. Figure 11 presents these two steps and how the Kripke jump $\mathcal{J}_m$ of the low-resolution reading combines them into one.

Whether or not Kripke's construction is characterized best with such high resolution depends on one's interests. For certain purposes, its standard, low resolution presentation is still advantageous. For example, in chapter 4 I will transfer Kripke's approach to a new area. I will develop and examine a model construction inspired by Kripke's, but using new methods and facing new challenges. For this, it proves advantageous to use the received low resolution framework because it is well understood and as such allows us to focus on difficulties specific to the new application.

For other purposes, again, it is worthwhile making use of the additional detail provided by my high resolution approach. For one, in the present section, the high resolution perspective has clarified what all variants of Kripke's construction have in common, their truth generator **T**, as well as how these variants differ, namely in their logic generator. For another, in chapter 8 I will develop a novel understanding of Kripke's construction. Here, too, high resolution will play a key role, as only it reveals natural, general principles exemplified by Kripke's construction, such as that a true conjunction is generated from its conjuncts.

I now turn to present *logic generators* that correspond to the non-classical, monotone evaluation schemes $m$ on the high-resolution picture.

## 3.5 STRONG KLEENE LOGIC

Recall the Strong Kleene evaluation scheme $\vDash_{sk}$, as defined, for example, in [Halbach, 2011, 15.10].[6]

6 As it is common in the literature, I deploy a slight strengthening of Kleene's original truth tables due to Albert Visser [1983].

**Definition 14** (Strong Kleene)**.** Let $\mathcal{L}$ be a first order language with '$\neg$','$\vee$' and '$\forall$' as its primitive logical symbols. Let $\mathfrak{M}$ be an $\mathcal{L}$-model that assigns to every $n$-place $\mathcal{L}$-relation symbol $R^n$ an extension $J^+(R^n)$ as well as an anti-extension $J^-(R^n)$. Let $\beta$ assign to every $\mathcal{L}$-variable an object of $\mathfrak{M}$'s domain $M$.

We define $\mathfrak{M} \models_{sk} \phi[\beta]$ by induction on the positive complexity of $\phi$.

$$
\begin{aligned}
\mathfrak{M} \models_{sk} R^n\overrightarrow{x_n}[\beta] &\text{ iff } \beta(\overrightarrow{x_n}) \in J^+(R^n)\\
\mathfrak{M} \models_{sk} \neg R^n\overrightarrow{x_n}[\beta] &\text{ iff } \beta(\overrightarrow{x_n}) \in J^-(R^n)\\
\mathfrak{M} \models_{sk} \neg\neg\phi[\beta] &\text{ iff } \mathfrak{M} \models_{sk} \phi[\beta]\\
\mathfrak{M} \models_{sk} \phi \vee \psi[\beta] &\text{ iff } \mathfrak{M} \models_{sk} \phi[\beta] \text{ or } \mathfrak{M} \models_{sk} \psi[\beta]\\
\mathfrak{M} \models_{sk} \neg(\phi \vee \psi)[\beta] &\text{ iff } \mathfrak{M} \models_{sk} \neg\phi \text{ and } \mathfrak{M} \models_{sk} \neg\psi[\beta]\\
\mathfrak{M} \models_{sk} \forall x\phi(x)[\beta] &\text{ iff for all } m \in M\, \mathfrak{M} \models_{sk} \phi(x)[\beta(x:m)]\\
\mathfrak{M} \models_{sk} \neg\forall x\phi(x)[\beta] &\text{ iff there is an } m \in M\, \mathfrak{M} \models_{sk} \neg\phi(x)[\beta(x:m)]
\end{aligned}
$$

In order to give the Strong Kleene Kripke jump $\mathcal{J}_{sk}$, we can recover the anti-extension of 'T' from its extension, in the following manner. Given a set of sentence X, let '$\neg X$' denote the set of all sentences $\phi$ such that $\neg\phi \in X$. If X comprises the sentences true under some interpretation of 'T', then $\neg X$ are the sentences false under it. $\neg X$ allows us to extract the anti-extension from the extension. Consequently, the Strong Kleene jump $\mathcal{J}_{sk}$ need not to work on pairs of sets but is well viewed as taking single sets of sentence codes X from which is extracted both positive and negative information.

$$\ulcorner\phi\urcorner \in \mathcal{J}_{sk}(X) \text{ iff } \mathfrak{N}(X, \neg X) \models_{sk} \phi \tag{12}$$

If we wish, as Kripke does, '$\neg Tx$' to hold of everything that is not a sentence, then we need to make one additional assumption. We let the set of sentence codes $\neg X$ contain not only all all codes $\ulcorner\phi\urcorner$ such that $\neg\phi\urcorner \in X$ but also all objects of the domain that do not encode a sentence. In what follows, I will tacitly assume that this trick is implemented.

Further, as indicated earlier I focus on the language of arithmetic $\mathcal{L}_a$ extended by a predicate symbol 'T' to the language of truth $\mathcal{L}_{ta}$. Its intended model are the standard numbers $\mathfrak{N}$ extended by some set of numbers as extension for 'T'. Consequently, we are in the convenient position that the language of our interest has a constant for every object of its intended domain. Thus, the quantifier clauses of the previous definition can be recast substitutionally, which allows us to drop the relativization to an assignment.

From the Strong Kleene jump (equation 12) we obtain the generator **JSK**, and on its basis the *low-resolution* characterization of Kripkean groundedness, relative to the Strong Kleene evaluation scheme. Figure 10 above shows one corresponding priority tree.

My present interest, however, is in the alternative, *high-resolution* understanding of semantic groundedness. It is groundedness through the combination of the truth generator **T** (p. 49) with some generator of monotone logic. In order to apply this schema to Kripke's Strong Kleene construction, I accordingly need to give a Strong Kleene logic generator. This is easily done: I turn the clauses of definition 14 into rules.

$$\frac{\phi}{(\phi \vee \psi)} \qquad \frac{\psi}{(\phi \vee \psi)} \tag{13}$$

$$\frac{\neg\phi \qquad \neg\psi}{\neg(\psi \vee \phi)} \qquad \frac{\phi}{\neg\neg\phi} \tag{14}$$

$$\frac{\psi(a) \qquad \psi(b) \qquad \ldots}{\forall x\big(\psi(x)\big)}\ a, b, \ldots \text{ are exactly the } \mathcal{L}\text{-constants} \tag{15}$$

$$\frac{\neg\psi(a)}{\neg\forall x\big(\psi(x)\big)}\ a \text{ is some such constant} \tag{16}$$

Recall, however, the Tarski generator **W** from section 2.3. The rules (13) to (16) are just those by which **W** was given there. Therefore we can let the Strong Kleene logic generator be just **W** and obtain a result analogous to fact 1 on p. 29. Note, however, that the language of arithmetic and truth $\mathcal{L}_{ta}$ has a term for every number $n$, its numeral $\overline{n}$. Therefore, we can skip its extension by constants.

**Lemma 4.** *Let $\Lambda$ be the set of $\mathcal{L}_a$-literals true in $\mathfrak{N}$. $\mathfrak{N}(X, \neg X) \models_{sk} \phi$ if and only if $\phi$ is **W**-grounded in $\Lambda$.*

*Proof.* Let me write '$\mathcal{T}(X)$' for the set $\{T\ulcorner\zeta\urcorner : \zeta \in X\} \cup \{\neg T\ulcorner\zeta\urcorner : \neg\zeta \in X\}$, and let $\Lambda$ be the set of $\mathcal{L}_a$-literals true in $\mathfrak{N}$. I show that $\mathfrak{N}(X, \neg X) \models_{sk} \phi$ if and only if $\phi$ is **W**-grounded in the sentences from $\Lambda \cup \mathcal{T}(X)$. For readability, I will equivocate between this set and its elements and call $\phi$ **W**-grounded in $\Lambda \cup \mathcal{T}(X)$.

Naturally, the lemma is proved by an induction on the positive complexity of $\phi$. At the base, let $\phi$ be a literal. If it does not contain 'T' then we have that $\phi$ holds in the model $\mathfrak{N}(X, \neg X)$ iff it is among the $\Lambda$, hence **W**-grounded in the $\Lambda$. So assume that $\phi$ is of the form $T\ulcorner\psi\urcorner$ or $\neg T\ulcorner\psi\urcorner$. We observe that

$$\mathfrak{N}(X, \neg X) \models_{sk} T\ulcorner\psi\urcorner \text{ iff } \psi \in X \text{ iff } T\ulcorner\psi\urcorner \in \{T\ulcorner\zeta\urcorner : \zeta \in X\}$$
$$\mathfrak{N}(X, \neg X) \models_{sk} \neg T\ulcorner\psi\urcorner \text{ iff } \psi \in \neg X \text{ iff } \neg T\ulcorner\psi\urcorner \in \{\neg T\ulcorner\zeta\urcorner : \neg\zeta \in X\}$$

Either way, $\phi$ is **W**-grounded in $\Lambda \cup \mathcal{T}(X)$.

At the induction step, let $\phi$ be the disjunction $\psi \vee \chi$ and assume that the lemma holds for both $\psi$ and $\chi$.

$$\begin{aligned}\mathfrak{N}(X,\neg X) \vDash_{sk} \psi \vee \chi \text{ iff } & \mathfrak{N}(X,\neg X) \vDash_{sk} \psi \text{ or } \mathfrak{N}(X,\neg X) \vDash_{sk} \chi \\ \text{iff } & \psi \text{ or } \chi \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X) \\ \text{iff } & \psi \vee \chi \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X)\end{aligned}$$

Now let $\phi$ be the negated disjunction $\neg(\psi \vee \chi)$.

$$\begin{aligned}\mathfrak{N}(X,\neg X) \vDash_{sk} \neg(\psi \vee \chi) \text{ iff } & \mathfrak{N}(X,\neg X) \vDash_{sk} \neg\psi \text{ and } \mathfrak{N}(X,\neg X) \vDash_{sk} \neg\chi \\ \overset{\text{I.H.}}{\text{iff}} & \neg\psi \text{ and } \neg\chi \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X) \\ \text{iff } & \neg(\psi \vee \chi) \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X)\end{aligned}$$

Finally, let $\phi$ be a quantified sentence $\forall x\psi(x)$.

$$\begin{aligned}\mathfrak{N}(X,\neg X) \vDash_{sk} \forall x\psi(x) \text{ iff } & \text{ for every } n \in \omega, \mathfrak{N}(X,\neg X) \vDash_{sk} \psi(\overline{n}) \\ \overset{\text{I.H.}}{\text{iff}} & \text{ for every } n \in \omega, \\ & \psi(\overline{n}) \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X) \\ \text{iff } & \forall x\psi(x) \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X)\end{aligned}$$

$$\begin{aligned}\mathfrak{N}(X,\neg X) \vDash_{sk} \neg\forall x\psi(x) \text{ iff } & \text{ for some } n \in \omega, \mathfrak{N}(X,\neg X) \vDash_{sk} \neg\psi(\overline{n}) \\ \overset{\text{I.H.}}{\text{iff}} & \text{ for some } n \in \omega, \\ & \psi(\overline{n}) \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X) \\ \text{iff } & \neg\forall x\psi(x) \ \mathbf{W}\text{-grounded in } \Lambda \cup \mathcal{T}(X)\end{aligned}$$

□

I now show, as announced in the previous section, that the sentences of Kripke's least Strong Kleene fixed point are precisely sentences **T**-**W** grounded in $\Lambda$.

**Lemma 5.** *For $\Lambda$ the set of $\mathcal{L}_a$-literals true in $\mathfrak{N}$ and every sentence $\phi$ of the extended language $\mathcal{L}_{ta}$, $\phi$ is in the least fixed point of Kripke's Strong Kleene jump just in case $\phi$ is **T**-**W**-grounded in the $\mathcal{L}_a$-literals true in $\mathfrak{N}$.*

*Proof.* As before, let $\Lambda$ the set of $\mathcal{L}_a$-literals true in $\mathfrak{N}$. Halbach [2011, 15.14] shows that a set of sentence codes $X$ is a $\mathcal{J}_{sk}$-fixed point if and only if it contains the (codes of the) $\Lambda$ and is closed under rules corresponding to T-Intro, ¬T-Intro (see my definition 13) and those from equations (13) to (16). In particular, the least $\mathcal{J}_{sk}$-fixed point is the least such set. □

Lemma 5 justifies the *high-resolution* understanding of Kripke's semantic groundedness based on Strong Kleene logic. It allows us to view the grounded sentences as generated from the arithmetical truths, by the combined application of the general truth generator **T** and the

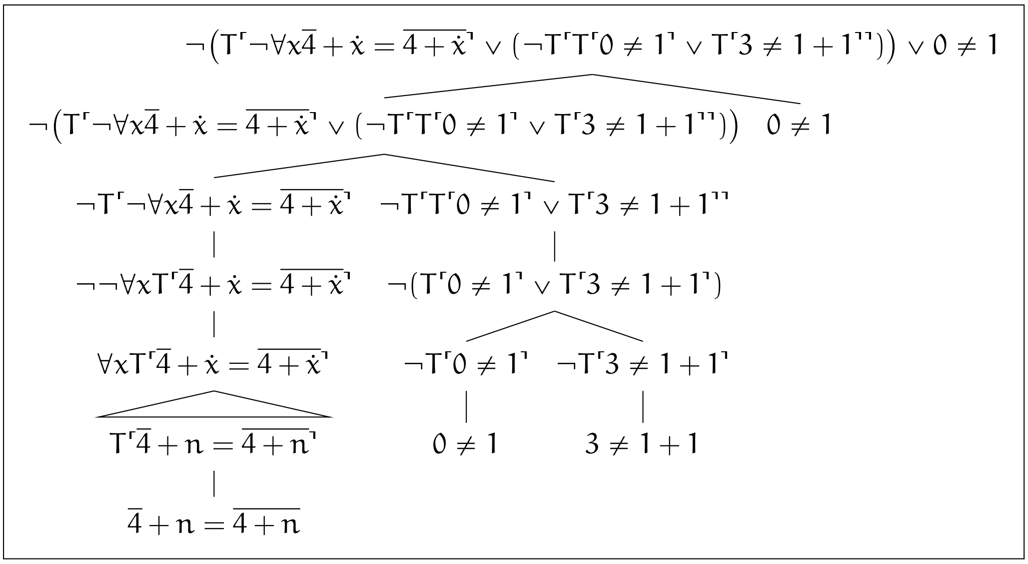


Figure 12: High resolution Kripke Groundedness: An Exemplary **T**-**W** Priority Tree

Strong Kleene logic generator **W**. To see the advantages of this high-resolution characterization of Strong Kleene groundedness, compare the **T**-**W**-priority tree in figure 12 with its low-resolution analogue 10. Note in particular how the former elucidates immediate dependencies such as that of $\neg(T\ulcorner 0 \neq 1\urcorner \vee T\ulcorner 3 \neq 1+1\urcorner)$ on $\neg T\ulcorner 0 \neq 1\urcorner$. To this extent, the high-resolution understanding of semantic groundedness developed in the previous section, paired with the general downwards perspective on groundedness from chapter 2, section 2.4, improves on Yablo's early characterization of dependence relations Yablo [1982], which took, e.g., $\neg(T\ulcorner 0 \neq 1\urcorner \vee T\ulcorner 3 \neq 1+1\urcorner)$ not to depend on $\neg T\ulcorner 0 = 1\urcorner$ but directly on $0 \neq 1$.

A variant of Kripke's theory that has gained some attention only recently is based on a *Weak* Kleene evaluation scheme [Feferman, 2008; Fujimoto, 2010]. As on the Strong Kleene scheme considered in the previous section, a relation symbol is assigned both an extension and an anti-extension. A literal of the form $\neg T\ulcorner\psi\urcorner$ is true not in the absence of $\psi$ from the extension of 'T' but only if $\psi$ is present in its anti-extension. The schemes differ in how complex sentences are treated. On the Weak Kleene approach, a complex sentence is true only if every constituent clause has a definite truth value. For example, the disjunction $\phi \vee \psi$ is true only if both disjuncts are true, $\phi$ is true and $\psi$ is false, or $\phi$ is false but $\psi$ is true. Accordingly, the Weak Kleene evaluations scheme is defined like the Strong Kleene scheme (definition 14) except that the clauses for negated disjunction $\neg(\phi \vee \psi)$ and negated universal quantification $\neg\forall x\phi(x)$ are extended by further conditions.

Modulo these alternations, however, the Weak Kleene logic generator is obtained analogously to how above the **W** has been obtained from the definition of Strong Kleene satisfaction. Further, just analogously to the proof of lemma 5, we can prove the following lemma.

**Lemma 6.** *Let $\Lambda$ be as before, and $\phi$ be any sentence of the extended language $\mathcal{L}_{\mathrm{ta}}$. Let* **K** *but the Weak Kleene logic generator. We have that $\phi$ is in the least fixed point of Kripke's* Weak *Kleene jump just in case $\phi$ is* ***T-K****-grounded in the sentences in $\Lambda$.*

## 3.6 SUPERVALUATION

I now turn to Kripkean theories of truth based on supervaluational logic. As is well known, it differs from the Strong Kleene approach in being not compositional. A disjunction may be super-true without either disjunct being so. To give a notorious example, even though of course Kripke's least supervaluational fixed points contain neither the liar sentence $\lambda$ not its negation, the disjunction $\lambda \vee \neg\lambda$ is in these fixed points.

This specific character of supervaluational Kripke groundedness justifies a more detailed exposition. Firstly, I will develop the standard, *low-resolution* presentation of Kripke's theory, although in a more general format than usual. Then, I will develop a *high-resolution* presentation in terms of the truth generator **T** and specifically supervaluational logic generators. Since supervaluational logic is not compositional, however, these logic generators will be of a different form than the generator **W** of the previous section.

I begin with the customary, low-resolution presentation of supervaluational Kripke theories. Recall that this is the reading based on Kripke's jump operator $\mathcal{J}_{\mathrm{m}}$, which turns truth in a model into a new model. The idea behind a supervaluational Kripke jump $\mathcal{J}_{\mathrm{sv}}$ is the following. Given a set of sentence codes X, we consider arbitrary extensions Y of X, each of which induces a classical model $\mathfrak{N}(Y)$, with Y interpreting 'T'. Then we use plain classical semantics to determine which sentences come out true in all of these models and add exactly these sentence (codes) to the interpretation X from which we started.[7]

$$\ulcorner\phi\urcorner \in \mathcal{J}_{\mathrm{sv}}(X) \text{ iff } \forall Y(X \subseteq Y \Rightarrow \mathfrak{N}(Y) \models \phi) \tag{17}$$

Of course, the more extensions Y we consider, the less agreement there is between the models $\mathfrak{N}(Y)$, and the less sentences enter the truth predicate. Usually, therefore, an *admissibility* condition is imposed on the range of extensions considered. Which interpretation Y is considered admissible, depends on the set X. Therefore, I will focus on the relation of some set Y *admissibly extending* a set X, in

7 I write doubly lined arrows '$\Rightarrow$' to signal that equations (17) and following are statements in my meta-language. The simple arrow '$\rightarrow$' is reserved for the object language $\mathcal{L}_{\mathrm{ta}}$.

symbols: $X \Subset Y$. For example, Burgess [1986] considers a supervaluational Kripke jump that requires not only Y to extend X, but also it not to contain any sentence whose negation is already found in X. Let '$\overline{X}$' denote the complement of X, and recall that $\neg X$ is the set $\{\phi : \neg\phi \in X\}$. Then, we can define Burgess' admissibility condition as follows.

$$X \Subset Y :\Leftrightarrow X \subseteq Y \subseteq \overline{\neg X} \tag{18}$$

Thus, Burgess restricts quantification to sets "sandwiched between" the given set X and those sentences whose negation does not occur in X. This choice of an admissibility condition gives rise to the following Kripke jump.

$$\ulcorner\phi\urcorner \in \mathcal{J}_{bs}(X) \text{ iff } \forall Y\big(X \subseteq Y \subseteq \overline{\neg X} \Rightarrow \mathfrak{N}(Y) \vDash \phi\big) \tag{19}$$

In the literature, various admissibility conditions $\Subset$ have been used. To give just one other example, Cantini [1990] works with a Kripkean theory based on the stronger admissibility condition of Y being a *consistent* extension of X.

$$\ulcorner\phi\urcorner \in \mathcal{J}_{cs}(X) \text{ iff } \forall Y\big(X \subseteq Y \,\&\, Y \textit{ consistent} \Rightarrow \mathfrak{N}(Y) \vDash \phi\big) \tag{20}$$

Other, stronger admissibility conditions are conceivable, too. In the following, I will reason schematically, for an arbitrary admissibility condition $\Subset$. In particular, I will not assume it to be definable but treat an admissibility condition as a class of candidate interpretations of 'T'. Thus, $\mathcal{J}_{bs}$ and $\mathcal{J}_{cs}$ are instances of the following general schema.

$$\ulcorner\phi\urcorner \in \mathcal{J}_{as}(X) \text{ iff } \forall Y(X \Subset Y \Rightarrow \mathfrak{N}(Y) \vDash \phi) \tag{21}$$

Given an admissibility condition $\Subset$ the corresponding Kripke jump $\mathcal{J}_{as}$ determines a way of generating sentences of the form $T\ulcorner\phi\urcorner$ from a given set of sentences X.

$$\textbf{JAS}\ \frac{X}{T\ulcorner\phi\urcorner} \text{ if } \phi \in \mathcal{J}_{as}(X) \tag{22}$$

We find that the least $\mathcal{J}_{as}$-fixed point comprises exactly the sentences grounded in nothing through this generator. This is not very surprising, as equation 22 does hardly more than rewriting the step from one stage of Kripke's construction to the next. Still, we have thus given a *low-resolution* reading of Kripke's supervaluational concept of grounded truth.

I now turn to the *high-resolution* understanding of Kripke's semantic groundedness. According to it, the sentences of Kripke's least fixed point based on the monotone evaluation scheme $m$ are grounded in the base truths through the combination of the truth generator **T** with a logic generator **M**. This **M** is the generator through which a sentence $\phi$ is grounded in the literals $T\ulcorner\zeta\urcorner$, $\neg T\ulcorner\xi\urcorner$ (plus the base language literals true in the base model), if and only if $\phi$ is $\mathfrak{m}$-true under the interpretation of 'T' by exactly those $\zeta, \neg\xi$. In the previous section, I gave such a generator **W** for Strong Kleene logic. Now, my goal is to identify a generator for supervaluation.

As noted above, supervaluational satisfaction schemes are not compositional. Consequently, the supervaluational generators **AS** will not be given by neat rules such as $\dfrac{\phi}{(\phi \vee \psi)}$. In order to generate a sentence $\phi$ from some other sentences X, we need to consider a range of admissible models of the language.[8] These admissible models, however, are all classical.

Recall from section 2.3 that taking the (classically) complete theory of a model $\mathfrak{N}(X)$ is to take the $\mathcal{L}_{ta}$-sentences **W**-grounded in the literals true in $\mathfrak{N}(X)$ (fact 1). This allows us to understand a supervaluational logic generator in terms of the **W**-generator from section 2.3. I write '$\mathcal{T}^\star(X)$' for the set $\{T\ulcorner\xi\urcorner : \xi \in X\} \cup \{T\ulcorner\xi\urcorner : \xi \notin X\}$. Note how $\mathcal{T}^\star(X)$ differs from $\mathcal{T}(X)$ of lemma 4: for any sentence $\xi$, $\mathcal{T}^\star(X)$ contains $T\ulcorner\xi\urcorner$ or $\neg T\ulcorner\xi\urcorner$. Consequently, $\mathcal{T}^\star(X)$ is ensured to contain either $T\ulcorner\xi\urcorner$ or its negation, for every sentence $\xi$ whether in X or not. This corresponds to the guiding idea of the supervaluational approach, to quantify over all *classical* extensions of a given model. After all, in a classical model $\mathfrak{N}(X)$, $\neg T\ulcorner\xi\urcorner$ is true iff $\xi \notin X$.

To define a supervaluational logic generator, I will make use of the relation which one set of literals $\mathcal{T}(X)$ bears to another $\mathcal{T}^\star(Y)$ if and only if Y *admissibly extends* X. This relation orders sets of literals according to how one interpretation of 'T' admissibly extends another. On this basis, the supervaluational generator **AS** is well viewed as a function of **W**-groundedness in admissible extension.

**Definition 15** (Supervaluational generators)**.** Given an admissibility condition $\Subset$ and a set of $\mathcal{L}_{ta}$-sentences X, let us say that $\phi$ is **AS**-generated from these sentences if and only if there is a set of sentences Y such that

1. $X = \mathcal{T}(Y)$ and

2. $\phi$ is strictly **W**-grounded in every $\mathcal{T}^\star(Z)$ such that $Y \Subset Z$.

Recall that something is said to be *strictly* $\beth$-grounded in some things if it takes at least one step to generate it from them. The second clause thus ensures that $\phi$, if **AS**-generated, is not itself of the form $T\ulcorner\psi\urcorner$ or $\neg T\ulcorner\psi\urcorner$.

By the first clause, supervaluational generators allow us to generate sentences only from literals of the form $T\ulcorner\psi\urcorner$ and $\neg T\ulcorner\psi\urcorner$. For example,

8 This is closely related to the fact that supervaluational consequence is not computably enumerable [Burgess, 1986; Fischer et al., 2014].

**BS** is the generator corresponding to Burgess' jump $\mathcal{J}_{bs}$ (equation 19). Based on fact 1, inspection of definition 15 leads to the following observation. It corresponds to lemma 4 of the previous section.

**Lemma 7.** *Let $\Subset$ be any admissibility condition and* ***AS*** *the corresponding generator according to definition 15. For any set* $X$*,* $\phi$ *is true in all* $\Subset$*-admissible extensions of* $\mathfrak{N}(X)$ *if and only if* $\phi$ *is a base language sentence true in* $\mathfrak{N}$*, or* ***T****-generated from* $X$ *or* ***AS****-generated from* $\mathcal{T}(X)$*.*

This fact allows us to finally show that Kripkean groundedness given by a supervaluational fixed point construction based on $\Subset$ is groundedness in true arithmetic through the corresponding generator **AS** combined with the Kripke truth generator **T**.

**Lemma 8.** *Let $\Subset$ be an admissibility condition,* ***AS*** *the corresponding generator, and* $\phi$ *an* $\mathcal{L}_{ta}$*-sentence. We have that* $\phi$ *is in the least fixed point of the Kripke jump based on* $\Subset$*, just in case* $\phi$ *is* ***T****-****AS****-grounded in the base language sentences* $\phi$ *true in* $\mathfrak{N}$*.*

*Proof.* Let $\Subset$ and $\phi$ be as required, and let '$\Theta$' denote the set of $\mathcal{L}_a$-sentences $\phi$ true in $\mathfrak{N}$.

Naturally, the left-to-right direction is shown by an induction on the low-resolution **JAS** rank of $\phi$. Its base holds trivially, because **JAS**-groundedness is groundedness in nothing. So let $\alpha + 1$ be the least ordinal such that $\phi \in I^{\alpha+1}_{JAS}$. Then $\phi \in \mathcal{J}_{as}(I^{\alpha}_{JAS})$. By lemma 7 we have that (†) $\phi$ is **T**-generated from some sentence $\psi \in I^{\alpha}_{JAS}$ or (‡) $\phi$ is **AS**-generated from $\mathcal{T}^{\star}(I^{\alpha}_{JAS})$.

If (†), then by the induction hypothesis $\phi$ is **T**-generated from some sentence $\psi$ that we know to be **T**-**AS**-grounded in $\Theta$, hence $\phi$ is, too. If (‡), then similarly $\phi$ is **AS**-generated from sentences that we already know to be **T**-**AS**-grounded in $\Theta$, hence $\phi$ is, too.

The right-to-left direction runs by an induction on the high-resolution **T**-**AS** rank, at whose base $\phi$ is ensured to be in the set $\Theta$, hence in the least $\mathcal{J}_{as}$-fixed point $I_{JAS}$. At the induction step, $\phi$ is either (†) **T**-generated or (‡) **AS**-generated from sentences **T**-**AS**-grounded in $\Theta$. Either way, by our induction hypothesis the sentences from which $\phi$ is generated are in the least fixed point of $\mathcal{J}_{as}$, hence their $I_{JAS}$-ranks have a least upper bound $\beta$. Consequently, they are all found in $I^{\beta}_{JAS}$.

If (†) then $\phi$ is **T**-generated from some sentences in $I^{\beta}_{JAS}$ or **AS**-generated from $\mathcal{T}^{\star}(I^{\beta}_{JAS})$. By lemma 7, therefore, $\phi \in \mathcal{J}_{as}(I^{\beta}_{JAS}) \in I_{JAS}$, as desired. □

## 3.7 CONCLUSION

In this chapter I applied my general theory of groundedness to its paradigm instance, Kripke's semantic groundedness. Having presented its standard formulation, I argued for a new, more finely grained characterization. The standard Kripke jump can be split into two steps, each of which corresponding to its own generator in the sense of chapter 2 (see figure 11). In a first step, from a given set of sentences $X$ we generate the literals of the form $T\ulcorner\phi\urcorner$, for $\phi$ in $X$, respectively $\neg T\ulcorner\phi\urcorner$, for $\neg\phi$ in $X$. This *truth* generator **T** is common to all variants of Kripke's construction. They differ in the second step which my analysis has brought to light, the logic generator. I showed that the resulting *high-resolution* characterization of semantic groundedness is co-extensional with the received low-resolution definition.

In the next chapter, I will turn to an instance of my general theory which unlike the case of sets or truth has not yet been sufficiently developed: groundedness models for type-free theories of concept-extensions, or classes.

# 4 GROUNDED CLASSES

## 4.1 MOTIVATION

Let $\mathcal{L}$ be some first-order language extended by a binary relation symbol '$\eta$'. The formula $x\eta y$ reads '$x$ is a member of $y$'. Consider the following schema of *naive comprehension*.

(C) For every $\mathcal{L}$-formula $\phi(x)$ with exactly the variable '$x$' occurring free in $\phi$,

$$\exists y \forall x (x \eta y \leftrightarrow \phi(x))$$

By Russell's paradox, (C) is inconsistent with classical first order logic. For, consider the $\mathcal{L}$-formula $x \not\eta x$. (C) requires there to be a class of everything that is not a member of itself. Instantiating for just this class, we find that it is a member of itself just in case it is not.

We regain consistency if we restrict the schema to formulae $\phi$ that do not contain '$\eta$'. Doing so, however, we undermine many interesting applications of class theory. For example, given a class $x$ we would like to have a class of the $y$ that are in some member of $x$. Thus, we would like to have the following instance of class comprehension.

$$\exists z \forall y \big(y \eta z \leftrightarrow \exists \nu (y \eta \nu \wedge z \eta x)\big) \tag{23}$$

Can we restrict comprehension in a more sophisticated manner, avoiding paradox while preserving its desirable instances? At this juncture, it is useful to look for inspiration elsewhere. Consider the case of extending a theory by a predicate symbol '$T$' that obeys Tarski's schema.

(T) For every sentence $\phi$,

$$T\ulcorner\phi\urcorner \leftrightarrow \phi$$

Of course, if we allow ourselves to substitute any sentence for $\phi$ then we face the paradoxes of truth, most prominently the Liar.

Tarski himself responded to the inconsistency of full schema (T) by restricting the schema to sentences which do not contain the truth predicate themselves. Adding this restricted schema to our base theory does not lead to paradox. In fact, the resulting theory has nice and natural models. However, this move also disallows many safe and indeed attractive instances of the full schema (T). For example, we would like to iterate applications of the truth predicate – but this is no longer possible.

Analogously to the case of class comprehension, therefore, banning the new relation symbol ensures consistency but comes at a cost. Fortunately, there is an alternative: Kripke's *groundedness* approach to truth, as discussed in the previous chapter.

Can we make a similar move in our present situation? Can we apply Kripke's method to single out the *grounded* instances of naive class comprehension? Extant literature gives reason to be hopeful. Most prominently, Penelope Maddy has carried out a Kripkean construction over set theory [Maddy, 1983, 2000].[1] The present chapter is intended as a general and systematic investigation into the prospects of grounded class theory. In the next section, I develop properties we would like such a theory to have. However, it is not guaranteed that these desiderata can all be satisfied; and maybe they need not all be. What follows are *prima facie* desirable features.

## 4.2 DESIDERATA FOR A THEORY OF GROUNDED CLASSES

Firstly, whichever way we approach a theory of grounded classes, we wish to answer Russell's paradox while allowing the membership relation to figure on the right-hand side of class comprehension. Thus, one desideratum is immediate. We want our theory to get us as much of comprehension as possible.

COMPREHENSION A class theory should contain many instances of class comprehension.

At this point, let me emphasize that although they pose analogous challenges, Tarski's schema and naive class comprehension differ in one respect. Whereas sentences are plugged into (T), the schema of comprehension takes open formulae; and these are universally quantified. As a result, one instance of (C) corresponds to many instances of (T). Comprehension for the formula $\phi(x)$ is grounded only if *for every* closed term $a$, the sentence $\phi(a)$ is grounded. In effect, as we will see, identifying grounded fragments of (C) is significantly more demanding than restricting the schema (T) to its grounded instances.

In order to motivate the second desideratum, allow me to ask: what do we need class theory for in the first place? After all, we already have a theory of *sets*, and it is both mathematically well developed and philosophically motivated. One way to argue that we also need a theory of classes is as follows[2]

There are two ways of collecting some things.[3] On the one hand, we collect some things by a sequence, possibly uncountable, of independent decisions whether a given object belongs to them or not – basically, by listing them. This *combinatorial* idea of collection underlies the theory of sets.

On the other hand, we collect some things by giving a condition which exactly they satisfy. This is the *definitional*, or *logical*, idea of

1 For an alternative approach, see Cantini [1996]. Mathematically, the theories also relate loosely to work by Feferman [1975a; 1975b] and Aczel [1980].

2 See [Maddy, 1983, §1] for the history of this line of thought.

3 Of course, from the Platonist viewpoint usually adopted, 'collection' strictly speaking is a metaphor. Much of the philosophy of set theory is devoted to explicating this metaphor, see e.g. Parsons [1977].

collection. For example, we may use the condition of being an ordinal number to collect, well, the ordinals. On pain of contradiction, there is no set of all the ordinals. Hence, in order to fully capture the definitional idea of collection, standard set theory needs to be supplemented by a theory of classes [Linnebo, 2006].[4] We would like to motivate our theory of grounded classes in this manner.

IDEA A class theory should stand to the *definitional* idea of collection as standard set theory stands to the combinatorial idea.

This desideratum is explicated naturally as follows. Defining conditions are closed under the logical connectives. Thus, we would like our classes to be closed under Boolean operations. For example, if according to our theory $x$ is not in the class of the $\phi$s, then $x$ must be in the class defined by the condition $\neg\phi$, in order for our theory to satisfy the desideratum. Further, there is a trivial condition (e.g. $x = x$) as well as one that nothing satisfies ($x \neq x$). Hence, our theory should have a universal and an empty class.

I turn to the next desideratum. By itself, the definitional idea leaves open when two conditions define the same collection. We may consider intensional identity criteria of different granularity.[5] My interest, however, is in those definitional collections the naive theory of which gave rise to Russell's paradox; and this notion of class, or concept-extension, is extensional. For example, the class of the ordinals is the class of the hereditarily transitive sets, since everything is an ordinal iff it is a hereditarily transitive set. Accordingly, our theory of grounded classes ought to make them extensional.[6]

EXTENSIONALITY A class theory should imply that the class of the $\phi$s is the class of the $\psi$s just in case: everything is a member of the class of the $\phi$s just in case it is a member of the class of the $\psi$s.

Finally, class talk is not peculiar to philosophers. Mathematicians speak of classes, too.[7] We would like our theory of classes to account for the usage of the notion in mathematics, at least for some of it.

How do working mathematicians use the notion of class? I will concentrate on two observations. On the one hand, the notion of class is used generally to speak of any collection which is not a set. In particular, different kinds of things are taken to form classes. Not merely sets, but numbers, graphs and categories. Consequently, our theory

4 To be explicit, I do not argue that standard set theory ought to be *replaced* by a theory of classes. Thus, the class theories developed below are not intended to play the role that, e.g., Quine's *New Foundations* is meant to fulfil.

5 Intensional theories of classes have been developed within the proof-theoretic programme of *explicit mathematics* [Feferman, 1975b, 1979; Jäger et al., 2001].

6 The set theoretic axiom of extensionality has been argued for on pragmatic, or external, grounds (Fraenkel et al. [1973], Maddy (1988, p. 483)). It seems to me that these arguments carry over to class theory.

7 See Parsons [1974] and Uzquiano [2003] for discussion of this point.

of grounded classes should be equally applicable to various areas. This intuitive thought must be rendered precise, however, as there are different senses in which a theory may be thought to be generally applicable. The relevant notion of applicability is this: we would like to be able to extend any given theory by classes grounded in it. It is in this specific sense of applicability that the following desideratum is to be understood.

BASE A class theory should be applicable to a variety of base theories.

On the other hand, mathematicians reason classically. Hence, we have the following desideratum.

CLASSICALITY A class theory should be closed under classical logic.

In this section, I have collected what we would, *prima facie*, a theory of grounded classes to be like. Next, I will explore how to develop such a theory.

It can be done in two ways. On the one hand, we may develop the theory directly, giving axioms or characterizing its intended model. Maddy followed this method [1983; 2000].[8] On the other hand, we may take a theory of grounded truth and translate it into the language of '$\eta$'.[9] The former, direct approach is arguably more natural. However, examining the latter, derivative method will illuminate challenges specific to class theory. Therefore, I will begin by exploring what can be done derivatively, and turn to the direct approach later (§4.5).

My presentation will be largely self-contained. As to notation, I will mostly follow Halbach [2011]. Deviation from or addition to his symbolism will be made explicit.

## 4.3 DERIVING GROUNDED CLASSES FROM GROUNDED TRUTH

In this section I examine theories of grounded classes derived from a given theory of grounded truth. The idea is this. We translate the language of class theory into the language of truth theory, roughly by translating $a\eta\ulcorner\phi\urcorner$ as $T\ulcorner\phi(a)\urcorner$.[10] Then, we endorse as our theory of grounded classes the set of sentences whose translations follow from our favourite theory of grounded truth.

To see how this works in detail, let us focus on the most popular theory of grounded truth, the theory of Kripke's least fixed point model based on Strong Kleene logic [Kripke, 1975]. Let $\mathcal{L}$ be the language of

8 The key technical idea is found already in Brady [1971]. See Hinnion and Libert [2003], §1, for a survey of this literature.

9 Work by Andrea Cantini can be viewed as being of this kind [1996, §§9-11].

10 Recall that $\ulcorner\phi\urcorner$ stands for $\phi$'s Gödel code or numeral, depending on the context.

first-order arithmetic plus '$\eta$'.[11] When dealing with paradox, caution is needed that may otherwise seem unnecessarily circumstantial. This concerns in particular the distinction of object- and meta-language. I will use letters from the beginning of the Roman alphabet ('$a$', '$b$' etc.), with decorations, as meta-linguistic variables for $\mathcal{L}$-terms and variables, and letters from the end of the Roman alphabet ('$x$', '$y$' etc.), decorations, as variables of the language $\mathcal{L}$. Fix a specific $\mathcal{L}$-variable $x_0$, and let *Fml* be the set of $\mathcal{L}$-formulae with $x_0$ as their single free variable.

In order to derive a class theory from Kripke's theory of truth, we translate an $\mathcal{L}$-sentence $\psi$ into the language of truth theory. To explain just how this is done will require me to go into some detail. Readers less formally inclined need not to follow me all the way; it suffices to keep in mind the basic idea that we translate $a\eta\ulcorner\phi\urcorner$ as $T\ulcorner(\phi)^\star(a)\urcorner$, for $(\phi)^\star$ the translation of $\phi$.

Usually, we define a translation by induction on syntactic complexity. The translation $(\cdot)^\star$, however, cannot be obtained in this manner, since in order to translate an atomic formula $a\eta\ulcorner\zeta \vee \xi\urcorner$ we must already have translated the complex formula $\zeta \vee \xi$. Towards an alternative definition of our translation, I propose the following notion of a formula's $\eta$-*rank*. Formulae of the base language have $\eta$-rank 0.

Also, formulae $a\eta b$ have $\eta$-rank 0 iff $b$ is a variable, or a closed term that is not a Gödel numeral $\ulcorner\phi\urcorner$. The $\eta$-rank of a formula '$a\,\eta\,\ulcorner\phi\urcorner$' is one greater than the $\eta$-rank of $\phi$. Complex formulae containing '$\eta$' inherit their $\eta$-rank from their immediate constituents. For example, the $\eta$-rank of $\phi \vee \psi$ is the $\eta$-rank of $\phi$ or $\psi$, whichever is greater; and the $\eta$-rank of $\exists x\phi$ is that of $\phi$. The fact that the code of $a\eta\ulcorner\phi\urcorner$ is strictly greater than that of $\phi$, ensures the relation "... is of lower $\eta$-rank than ..." to be well-founded on the $\mathcal{L}$-formulae. Thus, we can translate the language of class theory $\mathcal{L}$ into the language of truth theory.

A central role will be played by the syntactical operation *Sb* which takes a term $a$ and a formula $\phi \in Fml$, and outputs the substitution of $a$ for $x_0$ in $\phi$.[12] On the basis of our coding $\ulcorner\ldots\urcorner$, $Sb(a,\phi)$ is represented by an arithmetical formula $Sb^\bullet(x,y)$, such that first order arithmetic ('PA') proves $Sb^\bullet(\ulcorner a\urcorner,\ulcorner\phi\urcorner) = \ulcorner Sb(a,\phi)\urcorner$. Recall that I abbreviate by $\dot{x}$ a PA-representation of the function that maps a number $n$ to its numeral $\overline{n}$ (fn. 3 on p. 47). Quantification into the context $Sb^\bullet$ then is facilitated by quantification into this function, as in $\forall x\exists y\exists z\big(Sb^\bullet(\dot{x},\dot{y})\big) = z$. Occasionally, I will write $\ulcorner\phi(a)\urcorner$ for $Sb^\bullet(\ulcorner a\urcorner,\ulcorner\phi\urcorner)$.

11 For simplicity, I will assume '$\wedge$', '$\forall$' and '$\rightarrow$' to be defined in terms of the primitive symbols '$\neg$', '$\vee$' and '$\exists$'.

12 I assume that bound variables in $\phi$ are renamed if necessary.

**Definition 16.** Let $\mathcal{L}$ be the language of first order arithmetic $\mathcal{L}_a$ extended by '$\eta$' and let $\phi$ be an $\mathcal{L}$-formula. We define its translation $(\phi)^\star$ by induction on the $\eta$-rank of $\phi$.

If it is 0, then $\phi$ is a formula of the base language $\mathcal{L}_a$, or of the form $a\eta b$ and $b$ is not $\ulcorner\psi\urcorner$ for some formula $\psi$. If $\phi \in \mathcal{L}_a$ then we set $(\phi)^\star = \phi$. If $\phi$ is of the form $a\eta b$, and $b$ a variable, we set

$$(a\eta b)^\star = \begin{cases} \mathrm{T}Sb^\bullet(\ulcorner a\urcorner, \dot{b}) & \text{if } a \text{ is a closed term} \\ \mathrm{T}Sb^\bullet(\dot{a}, \dot{b}) & \text{if } a \text{ is a variable} \end{cases}$$

Finally, if $b$ is a closed term but does not denote the code of some formula, let $(\phi)^\star$ be $\mathrm{T}b$.

Now assume that the $\eta$-rank of $\phi$ is $n+1$, and that we have defined $(\zeta)^\star$ for formulae $\zeta$ of $\eta$-rank $\leqslant n$. At this point, inside of the induction on $\eta$-rank, we run an induction on the syntactic complexity of $\phi$. If $\phi$ is atomic, it is of the form $a\eta\ulcorner\zeta\urcorner$ for some formula $\zeta$ of $\eta$-rank $n$. We let

$$(a\eta\ulcorner\zeta\urcorner)^\star = \begin{cases} \mathrm{T}Sb^\bullet(\ulcorner a\urcorner, \ulcorner(\zeta)^{\star}\urcorner) & \text{if } a \text{ is a closed term} \\ \mathrm{T}Sb^\bullet(\dot{a}, \ulcorner(\zeta)^{\star}\urcorner) & \text{if } a \text{ is a variable} \end{cases}$$

Our induction hypothesis ensures $(\zeta)^\star$ to be defined. Now we set:

$$(\neg\phi)^\star = \neg(\phi)^\star$$

and proceed analogously for the other connectives and the quantifiers.

Using this translation we can define a theory in the language $\mathcal{L}$ as follows (recall § 3.5).

**Definition 17.** $\mathrm{HSK} := \{\phi : \mathfrak{N}(I_{sk}, \neg I_{sk}) \vDash_{\mathrm{SK}} (\phi)^\star\}$

I will speak of class theories using the following notation. The first letter 'H' indicates that we deal with a theory in a language containing '$\eta$'.[13] Then follows a code denoting the analogous truth theory. In the present case, 'SK' denotes the theory of the least Strong Kleene fixed point model.

In the following, I will examine HSK and test it against the desiderata of section 4.2. For this, I connect with notions due to Solomon Feferman.

Let $Cl(\ulcorner\phi\urcorner)$ be a meta-linguistic abbreviation of the formula $\forall y(y\eta\ulcorner\phi\urcorner \vee y\eta\ulcorner\neg\phi\urcorner)$[Feferman, 1991, p. 28] The property expressed by '$Cl$' will play a central role in the following. Note that HSK contains $Cl(\ulcorner\phi\urcorner)$

13 Recall that in the Greek alphabet, 'H' is a capital '$\eta$'.

just in case for every term $a$, the sentence $(\phi(a))^\star$ has a classical truth value in the Strong Kleene model $\mathfrak{N}(I_{sk}, \neg I_{sk})$ (recall definition 14). A sentence is classical in the least fixed point, however, just in case it satisfies Kripke's formal definition of groundedness. Therefore, if we seek a theory of grounded classes, then we ought to be interested into the condition *Cl*.

Moreover, if HSK contains $Cl(\ulcorner\phi\urcorner)$ then it contains $\forall x(x\eta\ulcorner\phi\urcorner \leftrightarrow \phi(x))$, and therefore the $\phi$-instance of comprehension. Due to this fact, I will say that a formula $\phi$ defines a class if $Cl(\ulcorner\phi\urcorner)$ holds in our theory, and will refer to *Cl* as the property of grounded class-hood.

Failure of grounded class-hood is identified fairly easily. $Cl(\ulcorner\phi\urcorner) \in$ HSK only if every for closed term $a$, the sentence $(\phi(a))^\star$ is grounded. Hence, the formula '$x_0\eta x_0$' fails to define a class since its instance '$\ulcorner x_0\eta x_0\urcorner\,\eta\,\ulcorner x_0\eta x_0\urcorner$' is translated as an ungrounded *truth-teller*. Similarly, comprehension does not hold for the Russell formula $x\not\eta x$. This is how HSK blocks the paradox from page 62.

It is good to know that the Russell formula does not define a class, but we would also like to know which formulae do so. More precisely, which formulae satisfy *Cl*, that is $\forall y(y\eta\ulcorner\phi\urcorner \vee y\eta\ulcorner\neg\phi\urcorner)$, over HSK? We can show that the theory contains all arithmetically definable classes, classes defined in terms of these, and so on. To render this precise, I introduce some terminology, again due to Feferman.[14]

**Definition 18.** Let $\phi$ and $\psi_0, \ldots, \psi_n$ be $\mathcal{L}$-formulae with exactly one free variable. Call $\phi$ *elementary in* the $\psi_i$ if (i) every atomic subformula in $\phi$ that contains '$\eta$' is of the form $a\eta\ulcorner\psi_i\urcorner$ for some $i \leqslant n$; and (ii) in $\phi$ only atomic subformulae are negated.

A formula $\phi$ is *elementary* simpliciter if there are some $\psi_i$ that $\phi$ is elementary in.

For example, $x\eta\ulcorner\psi\urcorner$ is elementary in $\psi$, as is $x\not\eta\ulcorner\psi\urcorner \vee \forall x\exists y(x = y + 1)$. The formula $\exists y(x\eta y)$, however, is not elementary, since it contains quantification into the range of $\eta$. This notion of elementarity allows us to give a sufficient condition on formulae $\phi$ for HSK to prove $Cl(\ulcorner\phi\urcorner)$.

**Proposition 6.** *For every* $\phi, \psi_0, ..., \psi_n \in$Fml *such that* $\phi$ *elementary in the* $\psi_i$, *if for every* $i \leqslant n$, $\mathrm{Cl}(\ulcorner\psi_i\urcorner) \in$ HSK *then*

$$\mathrm{Cl}(\ulcorner\phi\urcorner) \in \mathrm{HSK}$$

*Proof.* Recall that '$Cl(\ulcorner\phi\urcorner)$' abbreviates the $\mathcal{L}$-formula $\forall y(y\eta\hat{x}\phi \vee y\eta\hat{x}\neg\phi)$.

14 In the literature, 'elementary' usually applies to formulae of the base language. Feferman's concept is more general: elementary formulae may contain '$\eta$'. Indeed, they are closed under iteration of $\eta$. However, if $\phi$ is elementary in the $\psi_i$ then we know that in $\phi$, class talk is confined to atomic formula of the form $a\eta\ulcorner\psi_i\urcorner$. Thus, $\phi$ can be viewed as a base-linguistic function of atomic formulae $a\eta\ulcorner\psi_i\urcorner$: it is *elementary in* the $\psi_i$ [Feferman, 1975b].

Let $\phi$ be elementary in the $\psi_i$, and let for every $i$,

$$Cl(\ulcorner\psi_i\urcorner) \in \text{HSK}$$

.

We reason by induction on the syntactic complexity of $\phi$. If $\phi$ is atomic then it is either arithmetical, in which case we are done, or of the form $a\eta\ulcorner\psi_i\urcorner$, for some $i$. Since $Cl(\ulcorner\psi_i\urcorner)$, we have that for every $a$, the sentence $(\psi_i)^\star(a)$ has a classical truth value in Kripke's fixed point model for the language of truth. Consequently, so has $T\ulcorner(\psi_i)^\star(a)\urcorner$. Hence, either $T\ulcorner T\ulcorner(\psi_i)^\star(a)\urcorner\urcorner \in I_{sk}$ or $T\ulcorner\neg T\ulcorner(\psi_i)^{\star}\urcorner\urcorner \in I_{sk}$. Hence, by the fixed point character of $I_{sk}$ and our definition of HSK,

$$a\eta\ulcorner x\eta\ulcorner\psi_i\urcorner\urcorner \vee a\eta\ulcorner x\not\eta\ulcorner\psi_i\urcorner\urcorner$$

as desired.

Now let $\phi = \neg\zeta$. We know that $\zeta$ is an atomic formula. Again, if $\zeta$ is arithmetical then we are done. So assume that $\zeta$ is of the form $a\eta\ulcorner\psi_i\urcorner$ Since $Cl(\ulcorner\psi_i\urcorner) \in$HSK, we have that for every $a$, $T\ulcorner(\psi_i)^\star(a)\urcorner \in I_{sk}$ Hence, either $\neg\neg T\ulcorner(\psi_i)^\star(a)\urcorner \in I_{sk}$ or $\neg T\ulcorner(\psi_i)^\star(a)\urcorner \in I_{sk}$. Consequently, either $T\ulcorner\neg\neg T\ulcorner(\psi_i)^\star(a)\urcorner\urcorner \in I_{sk}$ or $T\ulcorner\neg T\ulcorner(\psi_i)^\star(a)\urcorner\urcorner \in I_{sk}$. Hence, $a\eta\ulcorner\phi\urcorner \vee a\eta\ulcorner\neg\phi\urcorner \in$HSK, as desired.

Now, let $\phi = \zeta \vee \xi$. Then, both $\zeta$ and $\xi$ must be elementary in the $\psi_i$, too, and our induction hypothesis ensures that $Cl(\ulcorner\zeta\urcorner) \in$ HSK and $Cl(\ulcorner\xi\urcorner) \in$ HSK. Consequently, for every term $a$ of $\mathcal{L}$, the sentences $(\zeta(a))^\star$ and $(\xi(a))^\star$ have a classical truth value in the Strong Kleene fixed point model. Consequently, the sentence $(\phi(a))^\star$ is ensured to have a classical truth value, too. Hence, $\forall x(x\eta\ulcorner\phi\urcorner \vee x\eta\ulcorner\neg\phi\urcorner) \in$ HSK.

Finally, assume that $\phi = \exists y\zeta(y, x_0)$ and that $\phi$ is elementary in the $\psi_i$. Then, so is $\zeta(a)$ for every $a$, and by our induction hypothesis, $Cl(\ulcorner\zeta(a)\urcorner) \in$ HSK. Hence, for every term $a$, every sentence $(\zeta(t,s))^\star$ has a classical truth value in the fixed point model. Hence, every sentence $(\exists y(\zeta(y,a)))^\star = (\phi(a))^\star$ is classical, and $T\ulcorner(\phi)^\star(a)\urcorner \vee T\ulcorner(\neg\phi)^\star(a)\urcorner$, i.e. $Cl(\ulcorner\phi\urcorner) \in$ HSK. $\square$

Proposition 6 proves useful. The notion of elementarity is purely syntactic. Thus, proposition 6 allows us to sidestep the non-classical semantics of Kripke's model construction and examine the class theory HSK directly. For one, every formula of the base language is classically equivalent to an elementary formula, and the base language fragment of HSK is closed under classical logic. Thus, HSK provides comprehension for arithmetical formulae. Consequently, it also proves comprehension for formulae elementary in arithmetical formulae. And so on. For another, there are elementary formulae every instance of which is true by logic, such as the formula '$x = x$'. For other elementary formulae, every instance is false by logic. Hence, HSK has a universal and an empty class. Furthermore, every formula of the form '$a = x$', for any closed term $a$, is trivially elementary and defines a class over HSK. Thus, the HSK classes are closed under the singleton operation.

Further, the syntactic property of elementarity is closed under the connectives. Consequently, the HSK classes are closed under the Boolean operations. For every $\phi$ and $\psi$, HSK firstly contains $Cl(\ulcorner\phi\urcorner)$ just in case it contains $Cl(\ulcorner\neg\phi\urcorner)$. Secondly, if HSK contains $Cl(\ulcorner\phi\urcorner)$ and $Cl(\ulcorner\psi\urcorner)$ then it contains $Cl(\ulcorner\phi \wedge \psi\urcorner)$ and $Cl(\ulcorner\phi \vee \psi\urcorner)$. Recall that for every formula $\phi$ such that $Cl(\ulcorner\phi\urcorner) \in$ HSK, HSK contains the corresponding instances of class comprehension. Together with the observations just made, this implies that the classes of HSK are closed under complement, union and intersection. On this basis, HSK can be viewed as capturing the definitional idea of collection, as we would like our theory of grounded classes to do.

How does HSK perform with respect to the other desiderata? We would like our class theory to be closed under classical logic. How does HSK perform in this respect? Badly. Of course, the model $\mathfrak{N}(I_{sk})$ is partial and the set of sentences true in it is not closed under classical logic. Consequently, HSK is not either. Hence, the theory HSK does not satisfy our desideratum of classicality.

Fortunately, another theory of grounded truth is closed under classical logic: Burgess' theory KFB [Burgess [2009] and [Halbach, 2011, §17]]. It axiomatizes the classical model $\mathfrak{N}(I_{sk})$, which we obtain from Kripke's partial model $\mathfrak{N}(I_{sk})$ by extending the anti-extension $\neg I_{sk}$ to the complement of the extension $I_{sk}$ (the *closed off* fixed point model). Further, KFB is a theory of grounded truth as it extends the well-known theory KF by a schema to the effect that '$\forall x\,(T\,(x) \rightarrow \phi\,(x))$' is proved whenever $\phi(x)$ satisfies the left-to-right direction of the KF axioms. In this precise sense, KFB axiomatizes the *least* predicate closed under the KF axioms. Since these correspond to the inductive clauses of Kripke's Strong Kleene fixed point construction, KFB is well viewed as an axiomatization of the *least* such fixed point.

The theory KFB has various properties which we would expect of a theory of grounded truth. For example, it proves the truth-teller sentences to be neither true nor false [Burgess, 2009, §14]. However, some features of KFB hardly square with semantic groundedness. Most prominently, the theory proves the Liar sentence, although not its truth.[15] On this basis, it may be challenged how faithful KFB is to the idea of semantic groundedness.

I do not wish to take a stance in this debate. However, if closing off the least fixed point is incompatible with groundedness, then the desideratum of classicality can hardly be met. In this chapter, I explore the prospects of grounded class theory, and will conclude that they are limited. Thus, I should first make a good case on behalf of the friend of grounded class theory. Therefore, I will examine what

15 KFB proves the theory KF+Cons [Halbach, 2011, §§17.2,17.3], which proves $\neg T\ulcorner\lambda\urcorner$, for PA $\vdash \lambda \leftrightarrow \neg T\ulcorner\lambda\urcorner$ [ibid., p. 217].

would be available to her if we assumed that closing off the least fixed point is compatible with the idea of groundedness. In sum, because I will argue that the prospects of grounded theories of classes are limited, it is fair to make an assumption on behalf of my opponent and concede the legitimacy of closing off. KFB axiomatizes the closed off least fixed point, and I will use it to obtain a theory of grounded classes.

Of course, I could also work with the complete theory of the closed off model $\mathfrak{N}(I_{sk})$. Unlike it, however, KFB is an axiomatic theory of truth. For some authors, the axiomatic approach to truth has advantages over the semantical approach. Although I do not claim that much, I wish to show how to obtain class theories from axiomatic as well as semantical theories of truth. HSK was based on a semantical theory of truth. Therefore, it is the axiomatic theory KFB from which I derive a theory of grounded classes HKFB, in the following manner.

**Definition 19.** $\mathrm{HKFB} \coloneqq \{\phi : \mathrm{KFB} \vdash (\phi)^{\star}\}$

HKFB has all desirable properties of HSK and excels in various other respects. To begin with, HKFB, unlike HSK, is closed under classical logic. It satisfies the desideratum of classicality. What fragment of naive comprehension does HKFB prove? As with HSK, the definition of '*Cl*' implies that for every $\phi$,

$$\mathrm{HKFB} \vdash Cl(\ulcorner\phi\urcorner) \rightarrow \forall x\big(x\eta\ulcorner\phi\urcorner \leftrightarrow \phi(x)\big) \tag{24}$$

Accordingly, the question again is: what formula does HKFB prove to have the property *Cl*? Above, I have found that in HSK, the set of formulae which define a class over HSK is closed under the connectives (proposition 6). The same holds for HKFB. Indeed, due to its classicality the theory proves the object-linguistic conditional.

**Proposition 7** (Cantini, 1996 9.7(ii))**.** *If $\phi$ is elementary in the $\psi_i$ then*

$$\mathrm{HKFB} \vdash \mathrm{Cl}(\ulcorner\psi_1\urcorner) \wedge \ldots \wedge \mathrm{Cl}(\ulcorner\psi_n\urcorner) \rightarrow \mathrm{Cl}(\ulcorner\phi\urcorner) \tag{25}$$

**Lemma 9** (Halbach 2011:§15.2, §17)**.**

$$\mathrm{KFB} \vdash \forall y(\neg(Ty \wedge T\dot{\neg}y))$$
$$\mathrm{KFB} \vdash \forall y(T\ulcorner\phi(\dot{y})\urcorner \rightarrow \phi(y))$$

*Proof of proposition 7.* Note that because $\phi$ is elementary in $\psi_i$ just in case $\neg\phi$ is elementary in $\psi_i$, the second conjunct in the consequent implies the first. Further, in view of lemma 9 it suffices to show

$$\mathrm{HKFB} \vdash Cl(\ulcorner\psi_0\urcorner) \wedge \ldots \wedge Cl(\ulcorner\psi_n\urcorner) \rightarrow \forall y(\phi(y) \rightarrow y\eta\ulcorner\phi\urcorner)$$

Observe that by classical logic (Boolean tautologies and quantifier negation), $\phi$ is equivalent to a formula $\phi'$ such that every negated subformula of $\phi'$ is a literal. That is, only atomic formula are negated in $\phi'$. This allows us to show the claim by induction on syntactic complexity of $\phi'$.

So assume $\phi'$ to be atomic. If $\phi'$ is an arithmetical formula then the claim follows directly from first two axioms of KFB, and if $\phi'$ translates as $Tx_0$, from axiom KFB12.

Now let $\phi$ be $\neg\chi$. We in fact know that $\chi = x_0\eta\ulcorner\psi_i\urcorner$ for some $i \leqslant n$. We need to show that

$$\mathrm{KFB} \vdash \forall y(T\ulcorner\psi_i(\dot{y})\urcorner \vee T\ulcorner\neg\psi_i(\dot{y})\urcorner) \to \forall y(\neg T\ulcorner\psi_i(\dot{y})\urcorner \to T\ulcorner\neg T\ulcorner\psi_i(\dot{y})\urcorner\urcorner)$$

We reason in KFB. Assume the antecedent and let $y$ be any object. If $T\ulcorner\psi_i(y)\urcorner$ then we are done. So assume $T\ulcorner\neg\psi_i(y)\urcorner$. The axiom KFB13 then allows us to conclude $T\ulcorner\neg T\ulcorner\psi_i(y)\urcorner\urcorner$, as desired.

For the other connectives and the quantifiers, the claim follows from the corresponding KFB axioms and the induction hypothesis. □

Schema 25 is proved already by weaker theories, in particular the $\mathcal{L}$-theory HKF+Cons derived from the theory of truth KF+Cons.[16] Thus, proposition 7 is not optimal from a proof-theoretic point of view. However, HKF+Cons cannot be viewed as a theory of *grounded* classes since, unlike KFB, it is not intended as a axiomatization of the least, but of all consistent Strong Kleene fixed points. From the philosophical perspective taken in this chapter, therefore, the theory HKFB is of particular interest.

As a corollary to proposition 7, HKFB itself proves the same closure of class-hood that we observed, meta-theoretically, for HSK (p. 70). To this extent, HKFB captures the definitional idea of collection and satisfies the corresponding desideratum (p. 64).

The theories considered so far, HSK and HKFB, extend first order arithmetic by a theory of classes. However, we would like our theory of grounded classes to be applicable to arbitrary base theories. To some extent, this poses a problem to the present, derivative approach. Grounded theories of truth are almost all developed over arithmetic. This restriction is useful, but fortunately not essential. Occasionally, other bases are considered. Recently, Kentaro Fujimoto examined the extension of ordinary set theory ZF by the truth axioms of KF [2012]. It can be strengthened to a theory of grounded truth ZF+KFB. Translating the language of set theory plus 'η' into the language of truth over set theory, we obtain a theory of grounded classes on top of ZF. We can show that its classes, too, are closed under elementary definition.[17]

16 See Cantini [1996, p. 7], and footnote 15 above. Cantini provides further information about a system mutually interpretable with HKF+Cons. For example, he shows that it interprets $\Sigma^1_1$-AC [Cantini, 1996, p. 66].

17 Based on the class-hood of elementary formulae, and generalizing a proof due to Feferman [1991], Fujimoto shows that his theory ZF+KF interprets iterations of NBG. For details, I refer the reader to Fujimoto's paper [2012].

So far, the derivative theory HKFB has performed well with respect to our desiderata. However, HKFB disappoints in one important respect: it does not satisfy the desideratum of extensionality, as I will show in the next section. I will have to go into some detail. The reader who accepts, if only for the sake of the argument, that extensionality poses a problem to the derivative approach, may well skip the following and continue with section 4.5 where I develop a new, direct approach to an extensional theory of classes.

## 4.4 DERIVATIVE THEORIES AND EXTENSIONALITY

I will begin by pointing out that even if distinct formulae $\phi$, $\psi$ define co-extensional classes, the theory HKFB is bound to contain $\ulcorner\phi\urcorner \neq \ulcorner\psi\urcorner$. Having noted this simple fact I will look more closely at what exactly is required for our theory of classes to satisfy the desideratum of extensionality. Based on this analysis, I will examine two routes that the friend of the derivative approach may take towards an extensional theory.

Firstly, I will discuss whether extensionality is at least achieved for the '='-free fragment of the language $\mathcal{L}$. This approach, however, puts undesirable limitations on our theory of classes.

Secondly, I pursue the thought that extensionality can be achieved by revising how we translate the language of classes into the language of truth. The idea is to translate $\ulcorner\phi\urcorner = \ulcorner\psi\urcorner$ as the statement that everything is a member of the class of the $\phi$s just in case it is a member of the class of the $\psi$s. As simple as this thought is, it will require some additional machinery to implement it. However, even if we make the necessary assumptions, we will find the resulting theory of classes not to satisfy the desideratum of extensionality. I conclude that instead of further elaborating on the derivative approach, we ought to develop a theory of grounded classes directly.

Consider any two equivalent arithmetical formulae, for example $\rho =$'$x_0 = \overline{2}$' and $\sigma =$'$x_0 = \overline{1} + \overline{1}$'. By proposition 7, HKFB contains

$$Cl(\ulcorner\rho\urcorner) \wedge Cl(\ulcorner\sigma\urcorner) \wedge \forall x(x\eta\ulcorner\rho\urcorner \leftrightarrow x\eta\ulcorner\sigma\urcorner) \tag{26}$$

Whichever reasonable way we choose to arithmetize syntax, $\rho$ and $\sigma$ are assigned distinct Gödel numbers. By arithmetic alone, therefore, HKFB also contains

$$\ulcorner\rho\urcorner \neq \ulcorner\sigma\urcorner \tag{27}$$

Thus, *prima facie* our theory says that there are co-extensional but distinct classes.[18]

18 This observation is a simple variant of known limitative results that apply to theories with comprehension for elementary formulae, or formulae of a similar syntactic property. See Gilmore [1974], Hinnion [1987], and (for a survey) Hinnion and Libert [2003].

In view of this basic fact, let us take a step back and ask what is required for our theory of classes to satisfy the desideratum of extensionality. Recall the desideratum: we would like our theory to imply that the class of the $\phi$s is the class of the $\psi$s just in case that everything is a member of the one if and only if it is a member of the other. On the present, derivative approach this means that we would like our class theory to say that the class of the $\phi$s is the class of the $\psi$s iff the underlying theory of truth proves $\forall x \big(T\ulcorner(\phi)^\star(x)\urcorner \leftrightarrow T\ulcorner(\psi)^\star(x)\urcorner\big)$. This schema induces an equivalence relation $E$ on the formulae in *Fml*. Thus, HKFB satisfies the desideratum of extensionality only if two formulae that stand in this relation $E$, define one and the same class. By Leibniz' law, the class of the $\phi$s is identical to the class of the $\psi$s only if one is indiscernible from the other. In the language $\mathcal{L}$, the class of the $\phi$s is denoted by the term $\ulcorner\phi\urcorner$. Therefore, HKFB satisfies extensionality only if for any two $E$-equivalent formulae $\phi$, $\psi$, it finds $\ulcorner\phi\urcorner$ and $\ulcorner\psi\urcorner$ indiscernible. The fact that HKFB contains both (26) and (27) shows that this is not so.

Maybe we have been too demanding. What (27) shows is that $\ulcorner\phi\urcorner$ and $\ulcorner\psi\urcorner$ are discernible *qua* codes. However, these distinct codes may still stand for the same class. All that matters is the following. If $\phi$ and $\psi$ stand in the relation $E$, $\ulcorner\phi\urcorner$ and $\ulcorner\psi\urcorner$ must not be *class-theoretically* discernible.

A natural way of rendering precise this thought is to consider the $\mathcal{L}$-fragment $\mathcal{L}^-$ without '$=$'. Then, we ask whether it holds that for every $\phi$ and $\psi$ of the original language $\mathcal{L}$, if $\phi$ bears $E$ to $\psi$ then $\ulcorner\phi\urcorner$ and $\ulcorner\psi\urcorner$ cannot be discerned within the fragment $\mathcal{L}^-$. More precisely, do we have that for every $\mathcal{L}^-$ formula $\zeta$, the theory HKFB contains $\zeta(\ulcorner\phi\urcorner) \leftrightarrow \zeta(\ulcorner\psi\urcorner)$? No. Let $\rho$ be as above, and let the number $n$ be its code. Since KFB proves $T\ulcorner\ulcorner\rho\urcorner = \overline{n}\urcorner$, our derived theory of classes contains $\ulcorner\rho\urcorner\,\eta\,\ulcorner x_0 = \overline{n}\urcorner$. However, KFB also proves $\neg T\ulcorner\ulcorner\sigma\urcorner = \overline{n}\urcorner$, for $\sigma$ as above. Therefore, HKFB contains $\ulcorner\sigma\urcorner \not\eta \ulcorner x_0 = \overline{n}\urcorner$; but '$y\,\eta\,\ulcorner x_0 = \overline{n}\urcorner$' is a formula of the '$=$'-free fragment $\mathcal{L}^-$. Hence, $\ulcorner\rho\urcorner$ and $\ulcorner\sigma\urcorner$ are not even indiscernible with respect to this restricted language.

It may be objected that although '$y\,\eta\,\ulcorner x_0 = \overline{n}\urcorner$' is an $\mathcal{L}^-$-formula, it contains the code of an equation. When asking for indiscernibility, the thought goes, we ought to not only focus on '$=$'-free formulae, but also disallow codes of formulae with '$=$'. However, the proposed notion of what makes a formula *class theoretic* is excessively restrictive: it gives up on classes defined by formulae of the base language. Our theory of grounded classes over arithmetic would not be able to speak of the arithmetical classes.

Fortunately, there is an alternative. The second route mentioned in the beginning of this section preserves class-definition in terms of '$=$'. Recall that its idea is to translate $\mathcal{L}$-formulae into the language of truth in a *smart* way. In order to implement this idea, I need to modify the setting of the derivate approach in two respects.

Initially, it may be thought that we can translate $\ulcorner\phi\urcorner = \ulcorner\psi\urcorner$ in such a way that our theory proves this base language sentence whenever $\phi$ and $\psi$ stand in the relation $E$. However, this would lead to a theory of classes that contradicts its own base theory. After all, for distinct formulae $\phi$ and $\psi$, $\ulcorner\phi\urcorner \neq \ulcorner\psi\urcorner$ is a theorem of rudimentary arithmetic.

In order to translate equations as statements of class-identity we need to disentangle the role of $\ulcorner\phi\urcorner$ as a number term and its role as standing for the formula $\phi$. A natural way of doing so is to speak of the class of the $\phi$s by a new term $\hat{x}\phi$, and no longer rely on its Gödel numeral $\ulcorner\phi\urcorner$. Formally, we extend the language $\mathcal{L}$ by a variable-binding, term-forming operator $\hat{\cdot}$ to the language $\mathcal{L}^\wedge$. We define the set of $\mathcal{L}^\wedge$-formulae and $\mathcal{L}^\wedge$-terms by simultaneous induction, such that $\hat{a}\phi$ is a term just in case $a$ is a variable and $\phi$ an $\mathcal{L}^\wedge$-formula. $\hat{a}\phi$ has precisely the free variables of $\phi$ but for $a$. I will refer to a term $\hat{a}\phi$ as an 'abstraction-term'.

In the remainder of this chapter, I will work with this extended language $\mathcal{L}^\wedge$. In addition to making the syntax of class theory more perspicuous, I have thus carried out the first of two changes that together will allow me to implement the *smart* way of translating class talk into truth talk.

The second modification concerns how the new, smart translation is defined. The notion of $\eta$-rank from the previous section is developed into the concept of a formula's $\eta$-*degree*. It is defined by an induction on $\phi$'s syntactic complexity, such that $a = b$ has $\eta$-degree 0 iff $a$ or $b$ is not an abstraction term $\hat{x}\phi$, and the $\eta$-degree of $\hat{x}\phi = \hat{x}\psi$ is one greater than that $\phi$ or $\psi$, whichever greater. The $\eta$-degree of a formula $a\eta b$ is defined just like its $\eta$-rank, and so is the $\eta$-degree of a syntactically complex formula. Since the term $\hat{x}\phi$ cannot occur in the formula $\phi$, the relation "... is of lower $\eta$-degree than ..." is well-founded.

Having extended the language by abstraction terms, and using the new concept of $\eta$-degree, we are now in a position to implement the smart translation of our language of class theory into the language of truth. Let $(\phi)^\dagger$ be defined by an induction on the $\eta$-degree of $\phi$. We proceed analogously to how we defined $(\phi)^\star$, except that now, $\hat{x}\zeta = \hat{x}\xi$ is translated as

$$\forall x \big(T\ulcorner(\zeta)^\dagger(x)\urcorner \leftrightarrow T\ulcorner(\xi)^\dagger(x)\urcorner\big) \tag{28}$$

Since the language of truth does not have abstraction terms, in other contexts $\hat{x}\zeta$ is represented by the code of the translation of $\zeta$. In particular, $\hat{x}\zeta\,\eta\,\hat{x}\xi$ is translated as $T\ulcorner(\xi)^\dagger\ulcorner(\zeta)^\dagger\urcorner\urcorner$. Shortly, we will find that this is a problem.

For the new translation $(\cdot)^\dagger$, the schema (28) defines over KFB a new equivalence relation $E_\dagger$. Let †HKFB be the theory of classes derived from the theory of truth KFB, through the smart translation $(\cdot)^\dagger$.

$$\dagger\mathrm{HKFB} \coloneqq \{\phi : \mathrm{KFB} \vdash (\phi)^\dagger\} \tag{29}$$

By definition, whenever $\phi$ and $\psi$ stand in the corresponding equivalence relation $E_\dagger$, $\dagger$HKFB contains $\hat{x}\phi = \hat{x}\zeta$ just in case it contains $\hat{x}\psi = \hat{x}\zeta$, for every $\zeta$. Thus, being smart about translating formulae $\hat{x}\phi = \hat{x}\psi$, we have come closer to our goal of an extensional theory of classes derived from KFB.

Have we succeeded? $\dagger$HKFB would satisfy the desideratum of extensionality if whenever $\phi$ and $\psi$ stand in the relation $E_\dagger$, we have that for every formula $\zeta$, $\dagger$HKFB contains $\zeta(\hat{x}\phi)$ just in case $\zeta(\hat{x}\psi)$. However, this is not the case. I will state the problem first, and then explain how it is rooted in the definition of $(\cdot)^\dagger$.

Let $\rho$, $\sigma$ be as above. Since KFB proves $\forall x\big(\rho(x) \leftrightarrow \sigma(x)\big)$, we have that $\dagger$HKFB contains $\hat{x}\rho = \hat{x}\sigma$. Now, let $m$ be the Gödel code of $(\rho)^\dagger$, such that PA$\vdash$ $\ulcorner(\rho)^{\dagger}\urcorner = \overline{m}$. We have that

$$\mathrm{KFB} \vdash T\ulcorner\ulcorner(\rho)^{\dagger}\urcorner = \overline{m}\urcorner \wedge \neg T\ulcorner\ulcorner(\sigma)^{\dagger}\urcorner = \overline{m}\urcorner \tag{30}$$

Consequently, our derived theory of classes is bound to contain the following.

$$\hat{x}\rho = \hat{x}\sigma \wedge \hat{x}\rho\,\eta\,\hat{x}(x = \overline{m}) \wedge \hat{x}\sigma \not\eta\, \hat{x}(x = \overline{m}) \tag{31}$$

Thus, it is not the case that formulae that stand in the equivalence relation $E_\dagger$ are indiscernible over the derived theory of classes $\dagger$HKFB. Therefore, although being based on a smart translation, the theory $\dagger$HKFB does not satisfy extensionality.

The reason is that if an abstraction term $\hat{x}\phi$ occurs on the left-hand side of '$\eta$', it is translated as the Gödel code of $(\phi)^\dagger$. More precisely, the formula $\hat{x}\phi\,\eta\,\hat{x}\zeta$ is translated as $T\ulcorner(\zeta)^{\dagger}(\ulcorner(\phi)^{\dagger}\urcorner)\urcorner$. However, this treatment of atomic formulae with '$\eta$' undoes what we have gained by being smart about '='. Since, all information is lost as to what other formulae bear $E_\dagger$ to $\phi$ when translating $\hat{x}\phi\,\eta\,\hat{x}\zeta$ as $T\ulcorner(\zeta)^{\dagger}(\ulcorner(\phi)^{\dagger}\urcorner)\urcorner$. As we have just seen, there are formulae involving '$\eta$' for which this information matters.

Although the translation $(\cdot)^\dagger$ is smarter than our original translation function $(\cdot)^\star$, it is not smart enough. In order to make the fact that $\phi$ and $\psi$ stand in the relation $E_\dagger$ ensure the indiscernibility of $\hat{x}\phi$ and $\hat{x}\psi$, not only $\hat{x}\phi = \hat{x}\psi$, but also $\hat{x}\phi\,\eta\,b$, for any $b$, needs to be translated in a manner that takes into account what other formulae are co-extensional with $\phi$.

One way of implementing this *smarter* approach is by translating $\hat{x}\phi\,\eta\,b$ and $\hat{x}\psi\,\eta\,b$ as the same formula of the language of truth, if $\phi$ and $\psi$ are co-extensional.[19] We may for example represent a formula $\phi$ by the lexicographically least $\mathcal{L}$-formula $\psi$ that bears $E_\dagger$ to $\phi$. That is, let $[\phi]_\dagger$ be the lexicographically least formula $\psi$ such that

$$\mathrm{KFB} \vdash \forall x\big(T\ulcorner(\phi)^{\dagger}(x)\urcorner \leftrightarrow T\ulcorner(\psi)^{\dagger}(x)\urcorner\big) \tag{32}$$

19 I thank Sam Roberts for this suggestion.

Thus, if distinct formulae $\phi$ and $\psi$ stand in the relation $E_\dagger$, they are both represented by the same formula $[\phi]_\dagger$. Using this representation, we can define a new translation $(\cdot)^\ddagger$ just like $(\cdot)^\dagger$ except that formulae $\hat{x}\phi\,\eta\,\hat{x}\psi$ are now translated as

$$T^{\ulcorner}(\psi)^{\ddagger\ulcorner}([\phi]_\dagger)^{\ddagger\urcorner\urcorner} \tag{33}$$

Let ‡HKFB be the $\mathcal{L}$-theory derived from KFB through this new, smarter translation $(\cdot)^\ddagger$. We undertook this further revision of our theory of classes in order to render $\hat{x}\phi$ and $\hat{x}\psi$ indiscernible, whenever $\phi$ and $\psi$ are co-extensional. To some extent, this has been achieved. Let $\phi$ and $\psi$ be distinct formulae that stand in the relation $E_\dagger$. Then, by equation 32, $[\phi]_\dagger = [\psi]_\dagger$. Trivially, we therefore have that KFB proves $T^{\ulcorner}(\zeta)^{\ddagger\ulcorner}([\phi]_\dagger)^{\ddagger\urcorner\urcorner} \leftrightarrow T^{\ulcorner}(\zeta)^{\ddagger\ulcorner}([\psi]_\dagger)^{\ddagger\urcorner\urcorner}$ for every $\zeta$. Consequently, $\mathrm{HKFB}^\ddagger$ contains $\forall z(\hat{x}\phi\,\eta\,z \leftrightarrow \hat{x}\psi\,\eta\,z)$, as desired. Thus, whenever $\phi$ and $\psi$ stand in the relation $E_\dagger$, $\hat{x}\phi$ and $\hat{x}\psi$ are indiscernible over $\mathrm{HKFB}^\ddagger$.

However, what we do not have is that $\hat{x}\phi$ and $\hat{x}\psi$ are thus indiscernible if $\phi$ and $\psi$ stand in the relation $E_\ddagger$ of our new, smarter translation $(\cdot)^\ddagger$ itself. That is, we do not have $\forall z(\hat{x}\phi\,\eta\,z \leftrightarrow \hat{x}\psi\,\eta\,z)$ for every $\phi$ and $\psi$ such that $\mathrm{KFB}\vdash \forall x\big(T^{\ulcorner}(\phi)^\ddagger(\dot{x})^{\urcorner} \leftrightarrow T^{\ulcorner}(\psi)^\ddagger(\dot{x})^{\urcorner}\big)$. The key change when moving from $(\cdot)^\dagger$ to $(\cdot)^\ddagger$ was that the new, smarter translation maps distinct formulae $\hat{x}\phi\,\eta\,\hat{x}\chi$, $\hat{x}\psi\,\eta\,\hat{x}\chi$ to one and the same formula $T^{\ulcorner}(\chi)^{\ddagger\ulcorner}([\phi]_\dagger)^{\ddagger\urcorner\urcorner} = T^{\ulcorner}(\chi)^{\ddagger\ulcorner}([\psi]_\dagger)^{\ddagger\urcorner\urcorner}$. For example, the formulae $\hat{x}\rho\,\eta\,\hat{x}(x = \overline{m})$, $\hat{x}\sigma\,\eta\,\hat{x}(x = \overline{m})$ from above are both mapped to $T^{\ulcorner\ulcorner}([\phi]_\dagger)^{\ddagger\urcorner} = \overline{m}^{\urcorner}$. More generally, we have that $\mathrm{HKFB}^\ddagger \vdash \forall y\big(\hat{x}\rho\,\eta\,y \leftrightarrow \hat{x}\sigma\,\eta\,y\big)$ such that $[\hat{x}\rho\,\eta\,y]_\ddagger = [\hat{x}\sigma\,\eta\,y]_\ddagger$; whereas, as we saw above, $\mathrm{HKFB}^\dagger \vdash \neg\forall y\big(\hat{x}\rho\,\eta\,y \leftrightarrow \hat{x}\sigma\,\eta\,y\big)$ such that $[\hat{x}\rho\,\eta\,y]_\dagger \neq [\hat{x}\sigma\,\eta\,y]_\dagger$. So far so good.

Now let o be the Gödel code of $([\hat{x}\rho\,\eta\,y]_\dagger)^\ddagger$. By reasoning very similar to how we established (31), we have that

$$\begin{aligned}\mathrm{HKFB}^\dagger \vdash \hat{y}(\hat{x}\rho\,\eta\,y) = \hat{y}(\hat{x}\sigma\,\eta\,y) \wedge \hat{y}(\hat{x}\rho\,\eta\,y)\eta\hat{x}(x = \overline{o})\\ \wedge\hat{y}(\hat{x}\sigma\,\eta\,y)\,\not\eta\,\hat{x}(x = \overline{o})\end{aligned} \tag{34}$$

In sum, formulae that stand in the relation $E_\dagger$ of the smart translation $(\cdot)^\dagger$ are indiscernible over the theory derived though the smarter translation $(\cdot)^\ddagger$. However, this smarter translation itself induces an equivalence relation $E_\ddagger$. As we have just seen, there are $E_\ddagger$-equivalent formulae that are not indiscernible over $\mathrm{HKFB}^\ddagger$. Therefore, even the further refinement of the derivative approach, based on the smarter translation $(\cdot)^\ddagger$, does not satisfy the desideratum of extensionality. Therefore, ‡HKFB does not satisfy extensionality, either. What would be needed is a translation t such that $\hat{x}\phi\,\eta\,\hat{x}\psi$ is translated as

$$T^{\ulcorner}(\psi)^{t\ulcorner}([\phi]_t)^{t\urcorner\urcorner} \tag{35}$$

Unfortunately, it is not obvious that such a mapping can be defined. In order for it to make sense to speak of $[\phi]_t$, t must already be defined not only for $\phi$, but for every other formula, too, including $\hat{x}\phi \, \eta \, \hat{x}\psi$ itself.

Independently of technical details, there is a philosophical reason not to pursue this route further. The more elaborate our translation, the less reason we have to think that the philosophical significance of our truth theory carries over to our derived theory of classes.[20] After all, syntactic translations do not generally preserve philosophical content. Moreover, the resources we invest in setting up a sophisticated translation we may as well use to develop a theory of classes directly. I will do so in the next section.

## 4.5 GROUNDED MEMBERSHIP AND GROUNDED IDENTITY

I will now develop a theory of grounded classes without the detour through truth theory characteristic of the derivative approach. My approach is semantical. I will define a model for the language with '$\eta$' and abstraction terms '$\hat{x}\phi$'. The basic idea is as follows. I will extend a given base model by a relation of class membership and a relation of class identity. These relations are defined inductively using jump operators that turn satisfaction in the given model into a new model. Together, these operators reach a least fixed point. In effect, I define *grounded* membership and *grounded* identity analogously to how Kripke defines a predicate of grounded truth (§ 3.3). Since doing so I enter a new area of groundedness, I will present my construction at *low* resolution, to render my proposal more accessible for readers familiar with the received characterization of semantic groundedness (see p. 50).

The model construction will combine elements of Penelope Maddy's theory [1983; 2000] as well as unpublished work by Hannes Leitgeb, and Leon Horsten and Øystein Linnebo. However, I will go beyond this extant work.

Maddy approaches a theory of grounded classes directly, and semantically. On the basis of set theory, she constructs a model for class theory using a monotone operator similar to my membership jump $\mathcal{H}$ below. Leitgeb, in an unpublished note from 2004, proceeds similarly. In the work of both authors, class identity is defined in terms of grounded membership. The class of the $\phi$s is the class of the $\psi$s if $\forall x(x\eta\hat{y}\phi \leftrightarrow x\eta\hat{y}\psi)$ holds in the least fixed point model. Effectively, class identity is dealt with as in the theory †HKFB of the previous section (p. 75). Consequently, Maddy's theory likewise fails to satisfy the desideratum of extensionality, as noted by herself [2000, p. 305]. The following is intended as one way of doing better.

In order to satisfy the desideratum of arbitrary bases, I will outline the construction for any first order base language $\mathcal{L}_a$, and any $\mathcal{L}_a$-structure $\mathfrak{M}$. For simplicity, I will assume that the base language

20 I return to the philosophical content of theories of grounded truth in chapter 8.

contains a constant for every object of the base domain.[21] Given the base model $\mathfrak{M}$, I proceed as follows.

Firstly, I extend the base domain $M$ by the set *Abs* of the closed abstraction terms. In the extended model, a closed term $\hat{x}\phi$ will denote itself. These terms will be the large pool of objects from which we will abstract the classes of our theory. It is useful to think of the terms as *proto*-classes, or class-candidates. For some terms $\hat{x}\phi$, the model below will validate $Cl(\hat{x}\phi)$ – the guiding question will be how many candidates are thus elected.[22]

Secondly, I add to the base model $\mathfrak{M}$ a membership relation $H$ and a relation of class identity $I$. In the extended model $\mathfrak{M}(I,H)$, the new relation symbol '$\eta$' will be interpreted by $H$.[23] Accordingly, $H$ relates objects of the full domain $M \cup Abs$ to proto-classes. $I$ extends plain identity on the base domain $M$ by a relation between proto-classes: $I \subseteq ID_M \cup Abs \times Abs$. Intuitively, $I$ extends identity in the base model by *class identity*. Accordingly, in the extended model $\mathfrak{M}(I,H)$, '=' will be interpreted by the relation $I$, such that, for example,

$$\mathfrak{M}(I,H) \models \hat{x}\phi = \hat{x}\psi \Leftrightarrow \langle \hat{x}\phi, \hat{x}\psi \rangle \in I \tag{36}$$

My goal is a specific model $\mathfrak{M}(I,H)$, a model for a theory of grounded classes. I will define a grounded membership relation $H$ and a grounded identity relation $I$, analogous to how Kripke defined a predicate of grounded truth.

My construction is based on two operators $\mathcal{I}$ and $\mathcal{H}$. Each takes one identity and one membership relation, but they differ in their output. $\mathcal{I}$, on the one hand, outputs an identity relation. I will speak of it as the 'identity jump'. $\mathcal{H}$, on the other hand, is a 'membership jump': it gives a membership relation.

There are various ways in which such jumps may be defined. I choose the *supervaluational* method, for two reasons. Firstly, doing so I explore an area not considered by Maddy [1983]. Secondly, I will eventually formulate a challenge to the friend of grounded classes. Therefore, I should first make a good case on her behalf. I will argue that the present, direct approach is unsatisfactory because it makes many natural candidates for class comprehension fail. More precisely, for many formulae $\phi$ that intuitively ought to define a class, the model does not validate either $a\eta\hat{x}\phi$ or $a\eta\hat{x}\neg\phi$ for every closed term $a$. Therefore, it is apposite to choose a semantics that maximizes the amount of sentences with classical truth value. Supervaluation fits this bill.

Recall the supervaluational variant of Kripke's least fixed point construction (§ 3.6 above). I the present, class-theoretic context, too, the

21 This is not the case if we work with the language of set theory. Here, we can either, as Fujimoto does, extend the language of set theory by new constants or work not with formulae, but with pairs of a formula and parameters. See Fujimoto [2012].

22 Recall that '$Cl(\hat{x}\phi)$' abbreviates the $\mathcal{L}$-formula $\forall y(y\eta\hat{x}\phi \vee y\eta\hat{x}\neg\phi)$.

23 Recall that 'H' here is the capital Greek letter eta.

basic idea is to consider a range of candidate interpretations of '=' and '$\eta$', determine which object o satisfies which formula $\phi$ in all these models and add $\langle o, \hat{x}\phi \rangle$ to the membership relation.[24] Analogously, we add $\langle \hat{x}\phi, \hat{x}\psi \rangle$ to the identity relation if $\phi$ and $\psi$ are coextensional in all models $\mathfrak{M}(J, K)$, for $J, K$ extending $I, H$. Since my interest is in relations J that are candidates for class identity, I will restrict my attention to *equivalence* relations that are *coherent* in the sense that if $\langle o, p \rangle \in J$ then for every formula $\phi$ $\langle \hat{x}\phi(x, \overline{o}), \hat{x}(x, \overline{p}) \rangle \in J$. Further, since my goal is a relation of class membership that respects class identity, I focus on pairs $J, K$ such that K respects J: for every $o, p$, if $\langle o, p \rangle \in J$ then for every q, $\langle o, q \rangle \in K$ if and only if $\langle p, q \rangle \in K$, and $\langle q, o \rangle \in K$ if and only if $\langle q, p \rangle \in K$. Below, this will allow me to show that grounded membership respects grounded identity, which in turn ensures the resulting theory to satisfy the desideratum of extensionality.

The more extensions we consider, the less pairs $\langle o, \hat{x}\phi \rangle$ will there be such that o satisfies $\phi$ in all of them (cf p. 55). Thus, the more extensions are considered, the less new information is added to the given relations of identity and membership. Hence, the more extensions are taken into account, the weaker our resulting theory will be. Therefore, usually further conditions are imposed on the range of extensions. The more restrictive such an admissibility condition, the more terms $\hat{x}\phi$ will be such that for every o either $\langle o, \hat{x}\phi \rangle$ or $\langle o, \hat{x}\neg\phi \rangle$ is added. Thus, which condition is chosen partly determines how many instances of class comprehension are satisfied.

Exploring the prospects of grounded class theory, I wish to test the best possible case for such a theory. For my model construction, I therefore choose the strongest admissibility condition available from the literature. In the variant of Kripke's jump operator due to Andrea Cantini, an extension is admissible if and only if it is consistent (recall equation (20) from p. 56) [Cantini, 1990, p. 250]. Its least fixed point exceeds all other supervaluational theories in the literature.[25] Accordingly, I will use jumps that quantify over consistent extensions only.[26]. In sum, a pair $J, K$ is an *admissible* extension of $I, H$ (in symbols: $I, H \Subset J, K$), if and only $I \subseteq J$, $H \subseteq K$, I is a *coherent* equivalence relation, K respects J, and they are both consistent.

I will now define the identity jump $\mathcal{I}$ and the membership jump $\mathcal{H}$. Firstly, the intuitive idea underlying the identity jump $\mathcal{I}$ is the following. $\mathcal{I}$ takes an identity relation I and a membership relation H

24 I use letters from the middle of the Roman alphabet ('n', 'o' etc.) as variables for objects of the extended domain $M \cup Abs$.

25 For example, Cantini's theory contains the sentences $\neg T\ulcorner\lambda\urcorner \vee T\ulcorner\neg\lambda\urcorner$, for liar sentences $\lambda$.

26 A membership relation H is consistent just in case there is no $\phi$ such that for any o, both $\langle o, \hat{x}\phi \rangle$ and $\langle o, \hat{x}\neg\phi \rangle$ are in H. An identity relation I is consistent if there is no $\phi$ such that for any o, both $\langle \hat{x}\phi, o \rangle \in I$ and $\langle \hat{x}\neg\phi, o \rangle \in I$

and identifies all pairs $\langle \hat{x}\phi, \hat{x}\psi \rangle$ such that $\phi$ and $\psi$ are co-extensional in all admissible extensions of the model $\mathfrak{N}(I, H)$.[27]

**Definition 20** (Identity Jump)**.**

$$\mathcal{I}(I, H) = \Big\{ \langle \hat{x}\phi, \hat{x}\psi \rangle : \forall J \forall K \Big( I, H \Subset J, K \Rightarrow \mathfrak{M}(J, K) \models \forall x \big( \phi(x) \leftrightarrow \psi(x) \big) \Big) \Big\}$$

I now turn to the *membership* jump $\mathcal{H}$. Its definition is based on the following idea. Given an identity relation I and a membership relation H, $\mathcal{H}$ outputs just the pairs $\langle o, \hat{x}\phi \rangle$ such that o satisfies the formulae $\phi$ in all models compatible with the identity and membership facts encoded in I and H. This intuitive idea is implemented by the supervaluational method which I have used already to obtain the identity jump $\mathcal{I}$. In order to identify the right pairs $\langle o, \hat{x}\phi \rangle$ we consider all admissible extensions of the pair $I, H$.[28]

**Definition 21** (Membership Jump)**.**

$$\mathcal{H}(I, H) = \big\{ \langle o, \hat{x}\phi \rangle : \forall K \forall J \big( I, H \Subset J, K \Rightarrow \mathfrak{M}(J, K) \models \phi(\overline{o}) \big) \big\}$$

I record some useful facts as to how $\mathcal{I}$ and $\mathcal{H}$ interact. Firstly, for any identity relation I and membership relation H, $\mathcal{I}(I, H)$ is an identity relation and $\mathcal{H}(I, H)$ is a membership relation. Secondly, if I and H are consistent, then so are $\mathcal{I}(I, H)$ and $\mathcal{H}(I, H)$. Finally, $\mathcal{I}(I, H)$ is an equivalence relation. Note, however, that neither is $\mathcal{H}(I, H)$ ensured to respect $\mathcal{I}(I, H)$, nor $\mathcal{I}(I, H)$ to be coherent.

Identity jump $\mathcal{I}$ and membership jump $\mathcal{H}$ together induce an operator on the pairs $I, H$. This operator is monotone with respect to the ordering of one pair of relations being extended pointwise by another. Therefore, it has a least fixed point $IH_\infty$.

I will denote the identity relation of the fixed point pair $IH_\infty$ by '$I_\infty$', and the membership relation by '$H_\infty$'. Note, however, that the interplay of the operators $\mathcal{I}$ and $\mathcal{H}$ is essential. It can be shown that $I_\infty$, which is obtained starting from the empty identity and the empty membership relation, is distinct from the least fixed point of $\mathcal{I}$, for the empty membership relation.

We can show the following key fact.

**Lemma 10.** $IH_\infty$ *is an admissible extension of itself.*

*Proof.* Firstly, of course, $I_\infty$ and $H_\infty$ extend themselves. Secondly, we have already observed that $I_\infty = \mathcal{I}(IH_\infty)$ is an equivalence relation (p. 81). Thirdly, we need to show that $H_\infty$ respects $I_\infty$, i.e.

27 In an unpublished manuscript, Leon Horsten and Øystein Linnebo use a similar jump to construct a model of Frege's Basic Law V. However, they keep the underlying second order logic predicative, as in Heck [1996]. In effect, their work corresponds to using $\mathcal{I}$ with a fixed membership relation H that captures satisfaction of base language formulae in the base model.

28 Recall that o has a name in our language $\mathcal{L}^\wedge$, that I will denote by '$\overline{o}$'.

A. for every $o, p \in M \cup Abs$, if $\langle o, p\rangle \in I_\infty$ then

  1. for all $q \in M \cup Abs$, $\langle q, o\rangle \in H_\infty$ iff $\langle q, p\rangle \in H_\infty$.
  2. for all $q \in M \cup Abs$, $\langle o, q\rangle \in H_\infty$ iff $\langle p, q\rangle \in H_\infty$.

Finally, we need to show the coherence of $I_\infty$, that is,

B. for every $o, p \in M \cup Abs$, if $\langle o, p\rangle \in I_\infty$ then for all formulae $\zeta(z, x)$ with exactly the free variables displayed, $\langle \hat{z}\zeta(z, \overline{o}), \hat{z}\zeta(z, \overline{p})\rangle \in I_\infty$

For simplicity, I focus on the case discussed in the main text, the fixed point over the natural numbers $\mathfrak{N}$.

(A1) If o or p is a base domain object, that we know not to be in the range of $H_\infty$, the claim is vacuously true. So let o be $\hat{x}\phi$ and p be $\hat{x}\psi$ for some $\phi, \psi$. By the definition of $\mathcal{H}$ we know that they have exactly one free variable x.

Assume that $\langle \hat{x}\phi, \hat{x}\psi\rangle \in I_\infty$, such that

$$\forall J \forall K \Big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \forall z\big(\phi(z) \Leftrightarrow \psi(z)\big)\Big) \tag{37}$$

Let q be any object. If $\langle q, \hat{x}\phi\rangle \in H_\infty$ then $\phi(\overline{q})$ holds at every admissible extension. By (37) and logic, we have $\mathfrak{N}(J, K) \models \psi(\overline{q})$ for every admissible extension $J, K$. Hence $\langle q, \hat{x}\psi\rangle \in H_\infty$, as desired. And analogously vice versa.

(A2) The claim is vacuously true unless q is a closed abstraction term $\hat{z}\zeta$. Since $IH_\infty$ is a fixed point, it suffices to show that

$$\forall J \forall K \big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \zeta(\overline{o})\big) \iff \forall J \forall K \big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \zeta(\overline{p})\big)$$

I show the left-to-right direction, the other is just analogous, swapping 'p' and 'o'. So assume the antecedent, and let $J, K$ be any admissible extension of $IH_\infty$. I show that $\mathfrak{N}(J, K) \models \zeta(\overline{p})$ by induction on the positive complexity of $\zeta$ [Halbach, 2011, definition 15.9].[29]

Firstly, $\zeta$ of the form $z = \overline{r}$ or $z \neq \overline{r}$, for some r, are taken care of by the *transitivity* of J together with our assumption that $\langle o, r\rangle \in I_\infty \subseteq J$. Secondly, let $\zeta$ be of the form $\hat{y}\xi(y, z) = \overline{r}$ or $\hat{y}\xi(y, z) \neq \overline{r}$. Since we assume that $\hat{y}\xi(y, \overline{o}) = \overline{r}$ holds in every admissible extension of $\mathfrak{N}(IH_\infty)$, the *coherence* of J ensures that $\langle \hat{y}\xi(y, \overline{p}), r\rangle \in J$ respectively $\langle \hat{y}\xi(y, \overline{p}), r\rangle \notin J$, and $\mathfrak{N}(J, K) \models \hat{y}\xi(y, \overline{p}) = \overline{r}$, as desired.

Thirdly, let $\zeta$ be of the form $z \eta a$ or $a \eta z$, respectively their negations. Then, $\mathfrak{N}(J, K) \models \zeta(\overline{p})$ follows from our assumption that $\mathfrak{N}(J, K) \models \zeta(\overline{o})$ and the fact that K respects J, which contains $\langle o, p\rangle$.

Finally, if $\zeta(\overline{o})$ is an $\eta$-literal such that $\overline{o}$ occurs within an abstraction term $b(x)$, we observe that the coherence of J ensures that $\langle b(\overline{o}), b(\overline{p})\rangle \in$

29 Careful examination shows that the attempt to prove the claim by induction on regular syntactic complexity breaks down at the induction step, at the clause for negations.

J.[30] Then, $\mathfrak{N}(J, K) \models \zeta(\overline{o})$ implies $\mathfrak{N}(J, K) \models \zeta(\overline{p})$ by the fact that K respects J.

At the induction step, we exploit the induction hypothesis. For example, let $\zeta(z)$ be of the form $\exists x\big(\xi(x, z)\big)$. Then for some object of the domain q, $\mathfrak{N}(J, K) \models \xi(\overline{q}, \overline{o})$. Now, $\xi(\overline{q}, z)$ is of lower complexity than $\zeta(z)$. Hence, our induction hypothesis ensures that $\mathfrak{N}(J, K) \models \xi(\overline{q}, \overline{p})$. Consequently, $\mathfrak{N}(J, K) \models \exists x\big(\xi(x, \overline{p})\big)$, as desired.

(B) Again, by the fixed point character of $\mathrm{IH}_\infty$, it suffices to show that

$$\forall J \forall K\Big(\mathrm{IH}_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \forall z\big(\zeta(z, \overline{o}) \leftrightarrow \zeta(z, \overline{p})\big)\Big)$$

So let $J, K$ be such that $\mathrm{IH}_\infty \Subset J, K$, and let q be any object of the domain. We show $\mathfrak{N}(J, K) \models \zeta(\overline{q}, \overline{o}) \leftrightarrow \zeta(\overline{q}, \overline{p})$ by induction on the complexity of $\zeta$. Recall that by the definition of the identity jump $\zeta$ is ensured to have exactly two free variables. I confine myself to the left-to-right direction as again, the other direction is just analogous.

We reason much like in the case of (A2). If $\zeta(\overline{q}, \overline{o})$ is of the form $\overline{q} = \overline{o}$ or $\hat{x}\xi(x, \overline{q}) = \overline{o}$ for some $\xi$, the claim follows from the transitivity of J. If it is of the form $\hat{y}\xi(y, \overline{o}) = \overline{q}$ we recall that J is coherent and contains $\langle o, p\rangle$. Finally, for atomic formulae containing $\eta$ we note that K respects the coherent J, as before. The induction step is taken care of by the induction hypothesis and logic. For example, if $\zeta(\overline{q}, \overline{o})$ is of the form $\exists x \xi(x, \overline{q}, \overline{o})$ we reason as we did at the end of the argument for (A2). □

I now examine what theory of classes this model construction provides. For a fair comparison with the theories of the previous sections, I focus on first-order arithmetic as our base theory. Thus, we extend the standard model of arithmetic $\mathfrak{N}$ by the least fixed point pair of relations $\mathrm{IH}_\infty$, obtained on the basis of arithmetic. The complete theory of this model I call 'HC', since the construction is based on the admissibility condition of *consistency*.

**Definition 22.**

$$\mathrm{HC} := \{\phi : \mathfrak{N}(\mathrm{IH}_\infty) \models \phi\}$$

I examine HC against the desiderata from section 4.2. Firstly, $\mathfrak{N}(\mathrm{IH}_\infty)$ is a classical model. Hence, HC meets the desideratum of classicality. So did the derivative theory HKFB considered above. Unlike in HKFB, however, every classically tautological formula defines a class over HC. In this precise sense, HC may be viewed as being more classical than HKFB. Of course, this additional degree of classicality is paid for. For example, it is not the case that $x\eta\hat{y}\phi$ or $x\eta\hat{y}\psi$ whenever $x\eta\hat{y}(\phi \vee \psi)$. This fact is due to the choice of supervaluational

30 Our definition of the identity jump operator $\mathcal{I}$ ensures $\zeta$ to have exactly one free variable, which in this case implies that $b(x)$, too, has just the free variable displayed.

operators, and corresponds to the failure of compositionality in supervaluational truth theory. However, in the present class-theoretic context I consider it less problematic. We are interested not in single sentences of the form $x\eta\hat{x}(\phi \vee \psi)$, but in formulae $\phi \vee \psi$ of which we know that they define classes. And in this respect, a supervaluational class theory is not inferior to a closed-off Strong Kleene theory such as HKFB – neither proves downwards closure of classes under the connectives, e.g. $Cl\big(\hat{x}(\phi \vee \psi)\big) \rightarrow Cl(\hat{x}\phi) \vee Cl(\hat{x}\psi)$.

Secondly, class theories should stand to the *definitional* idea of collection as standard set theory stands to the *combinatorial* idea. How does the theory HC do with respect to this desideratum? A natural sharpening of the definitional idea was that classes make up a Boolean algebra. Brief reflection on the fixed point character of the model $\mathfrak{N}(IH_\infty)$ and its classicality shows that the theory HC is closed under complement, union, intersection and iteration of membership.

**Proposition 8.**

$$\begin{aligned}\mathfrak{N}(IH_\infty) \models \forall x \forall y \Big(Cl(x) \wedge Cl(y) \rightarrow & \exists z\big(Cl(z) \wedge \forall w(w\eta z \leftrightarrow r \not\eta x)\big) \\ & \wedge \exists z\big(Cl(z) \wedge \forall w(w\eta z \leftrightarrow r\, \eta\, x \vee w\eta y)\big) \\ & \wedge \exists z\big(Cl(z) \wedge \forall w(w\eta z \leftrightarrow r\, \eta\, x \wedge w\eta y)\big) \\ & \wedge Cl\big(\hat{z}(z\eta x)\big)\Big)\end{aligned}$$

It is easy to see that every base language formula $\phi$ defines a class.[31] More precisely, for every $\mathcal{L}_a$-formula $\phi$ with a single free variable we have that

$$\mathfrak{N}(IH_\infty) \models Cl(\hat{x}\phi) \tag{38}$$

Thus, HC is a theory of classes based on first-order arithmetic and closed under natural operations. This makes HC a good candidate for a formal theory of definitional collections.

Thirdly, the desideratum of extensionality was met by none of the derivative theories. HC performs considerably better in this respect. On the one hand, since $I_\infty$ is a fixed point of the identity jump $\mathcal{I}$, two terms $\hat{x}\phi$ and $\hat{x}\psi$ stand in the relation $I_\infty$ just in case they are co-extensional in $\mathfrak{N}(IH_\infty)$.

**Proposition 9.** *For all* $o, p \in$Abs,

$$\mathfrak{N}(IH_\infty) \models \overline{o} = \overline{p} \leftrightarrow \forall z(z\eta\overline{o} \leftrightarrow z\eta\overline{p})$$

*Proof.* The left-to-right direction of the claim follows directly from the fact that $H_\infty$ respects $I_\infty$ (lemma 10). For the right-to-left direction, assume that for every $q$, $\langle q, \hat{x}\phi\rangle \in H_\infty$ iff $\langle q, \hat{x}\psi\rangle \in H_\infty$. Hence,

31 For $o$ $\in$*Abs*, $\mathfrak{N}(IH_\infty) \models \neg\phi(\overline{o})$ such that $\langle o, \hat{x}\neg\phi\rangle \in H_\infty$; and for every $o \in \omega$, $\mathfrak{N} \models \phi(\overline{o}) \vee \neg\phi(\overline{o})$ such that $\langle o, \hat{x}\phi\rangle$ or $\langle o, \hat{x}\neg\phi\rangle$ enters the membership relation at the very first stage.

$\forall J \forall K\big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \phi(\overline{q})\big)$ iff $\forall J \forall K\big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \psi(\overline{q})\big)$ By logic, $\forall J \forall K\Big(IH_\infty \Subset J, K \Rightarrow \big(\mathfrak{N}(J,K) \vDash \phi(\overline{q}) \Leftrightarrow \mathfrak{N}(J,K) \vDash \psi(\overline{q})\Big)$. Since this holds for every q, $\forall J \forall K\Big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \forall z\big(\phi(z) \leftrightarrow \psi(z)\big)\Big)$. Hence, $\langle \hat{x}\phi, \hat{x}\psi \rangle \in I_\infty$, as desired.

□

On the other hand, every two class terms that stand in the relation $I_\infty$ are also ensured to be *indiscernible* in the model $\mathfrak{N}(IH_\infty)$.[32] I abbreviate a list of variables $x_0, \ldots, x_n$ as '$\overrightarrow{x_n}$'.

**Proposition 10.** *For every* $o, p \in M \cup Abs$, *if* $\langle o, p \rangle \in I_\infty$ *then we have that for every* $\mathcal{L}^\wedge$*-formula* $\phi(\overrightarrow{x_{n+1}})$,

$$\mathfrak{N}(IH_\infty) \vDash \forall \overrightarrow{x_n}\big(\phi(\overline{o}, \overrightarrow{x_n}) \leftrightarrow \phi(\overline{p}, \overrightarrow{x_n})\big)$$

*Proof.* Let $\langle o, p \rangle \in I_\infty$. A basic theorem of model theory says that if two objects are indiscernible in a first-order model $\mathfrak{M}$ in terms of the primitive relation symbols of the $\mathfrak{M}$ signature, then they are indiscernible in $\mathfrak{M}$ with respect to every formula $\phi$ of this language [Ketland, 2011, lemma 3.5]. It is proved by induction on the complexity of $\phi$.

The following is a modification of that standard proof, for the nonstandard language $\mathcal{L}^\wedge$ with its abstraction terms. It has two primitive relation symbols, '=' and '$\eta$'. Therefore, we have to show that every two o and p that stand in the relation $I_\infty$ are indiscernible in terms of '=' and '$\eta$'. Since $\overline{o}$ and $\overline{p}$, however, may occur within open abstraction terms, six cases need to be distinguished.

I. $\mathfrak{N}(IH_\infty) \vDash \forall \overrightarrow{x_n}\big(a(\overrightarrow{x_n})\,\eta\,\overline{o} \leftrightarrow a(\overrightarrow{x_n})\,\eta\,\overline{p}\big)$, for every $a(\overrightarrow{x_n})$.

II. $\mathfrak{N}(IH_\infty) \vDash \forall \overrightarrow{x_n}\big(\overline{o}\,\eta\,a(\overrightarrow{x_n}) \leftrightarrow \overline{p}\,\eta\,a(\overrightarrow{x_n})\big)$, for every $a(\overrightarrow{x_n})$.

(I) and (II) follow directly from the fact that $H_\infty$ respects $I_\infty$ (lemma 10). From the coherence of $I_\infty$ we know that it contains $\langle b(\overline{o}, \overline{q_0}, \ldots, \overline{q_n}), b(\overline{p}, \overline{q_0}, \ldots, \overline{q_n}) \rangle$ for every term b with $n+1$ free variables and every sequence of objects $q_0, \ldots, q_n$ with their canonical $\mathcal{L}^\wedge$-terms $\overline{q_0}, \ldots, \overline{q_n}$. From the fact that $H_\infty$ respects $I_\infty$ it follows that

III. $\mathfrak{N}(IH_\infty) \vDash \forall \overrightarrow{x_n}\big(a(\overrightarrow{x_n})\,\eta\,b(\overline{o}, \overrightarrow{x_n}) \leftrightarrow a(\overrightarrow{x_n})\,\eta\,b(\overline{p}, \overrightarrow{x_n})\big)$, for every $a(\overrightarrow{x_n})$ and $b(\overrightarrow{x_{n+1}})$.

IV. $\mathfrak{N}(IH_\infty) \vDash \forall \overrightarrow{x_n}\big(b(\overline{o}, \overrightarrow{x_n})\,\eta\,a(\overrightarrow{x_n}) \leftrightarrow b(\overline{p}, \overrightarrow{x_n})\,\eta\,a(\overrightarrow{x_n})\big)$, for every $a(\overrightarrow{x_n})$ and $b(\overrightarrow{x_{n+1}})$.

32 Proposition 10 strengthens an early result by Ross Brady, who shows, modulo notation, that the schema $\forall y(y\eta\hat{x}\phi \leftrightarrow y\eta\hat{x}\psi)$ over his theory defines a *congruent* relation of co-extensionality. Unlike the language $\mathcal{L}$ of the theory HC, however, Brady's language does not have identity Brady [1971].

Since $I_\infty$ is an equivalence relation, in particular transitive, we have that

V. $\mathfrak{N}(\mathrm{IH}_\infty) \models \forall \overrightarrow{x_n}\big(a(\overrightarrow{x_n}) = \overline{o} \leftrightarrow a(\overrightarrow{x_n}) = \overline{p}\big)$, for all terms $a(\overrightarrow{x_n})$.

Finally,

VI. $\mathfrak{N}(\mathrm{IH}_\infty) \models \forall \overrightarrow{x_n}\big(a(\overrightarrow{x_n}) = b(\overline{o}, \overrightarrow{x_n}) \leftrightarrow a(\overrightarrow{x_n}) = b(\overline{p}, \overrightarrow{x_n})\big)$, for all terms $a(\overrightarrow{x_n})$ and $b(\overrightarrow{x_{n+1}})$.

is true because firstly $I_\infty$ is coherent, such that for every sequence of objects $q_0, \ldots, q_n$ we have that $\langle b(\overline{o}, \overline{q_0}, \ldots, \overline{q_n}), b(\overline{p}, \overline{q_0}, \ldots, \overline{q_n})\rangle \in I_\infty$; secondly, the transitivity of $I_\infty$ ensures that for every $r$,

$$\langle b(\overline{o}, \overline{q_0}, \ldots, \overline{q_n}), r\rangle \in I_\infty \text{ iff } \langle b(\overline{p}, \overline{q_0}, \ldots, \overline{q_n}), r\rangle \in I_\infty$$

, as desired.

Based on (I) to (VI), we show by an ordinary induction on the syntactic complexity of $\phi$ that

$$\mathfrak{N}(\mathrm{IH}_\infty) \models \forall \overrightarrow{x_n}\big(\phi(\overline{o}, \overrightarrow{x_n}) \leftrightarrow \phi(\overline{p}, \overrightarrow{x_n})\big)$$

thus completing the proof. □

Result 10 is highly desirable, and distinguishes HC from all other theories considered in this chapter (see §4.4 above).

So far, the theory HC of the fixed point model $\mathfrak{N}(\mathrm{IH}_\infty)$ has performed well. I now turn to the desideratum of comprehension. How much of the comprehension schema does HC contain? As it has been the case with the derivative theories of the previous sections, HC contains comprehension for a formula $\phi$ just in case it contains $Cl(\hat{x}\phi)$. Above, we have seen that every base language formula $\phi$ defines a class. However, the goal of a grounded theory of classes is to recover as much comprehension as possible for formulae that contain '$\eta$'.

The derivative theory HKFB proved class-hood, and thus comprehension, for every *elementary* formula. Unfortunately, this positive result does not carry over to the present, direct approach. Over the class theory HC, elementarity no longer suffices for class-hood.

To see this, consider any formula $\phi$ elementary in the $\psi_i$, and assume that HC proves these $\psi_i$ to define classes. In the old setting, this sufficed for $\phi$, too, to define a class, even if $\phi$ contains an atomic formula of the form $\ulcorner\zeta\urcorner = a$, for some $\zeta$ not among the $\psi_i$. It only mattered which terms occur in the range of $\eta$. Formulae such as '$\ulcorner\zeta\urcorner = x_0$' did not incur presuppositions about $\zeta$. However, the very point of the present model construction was a more sophisticated treatment of identity statements. As a consequence, however, elementarity as in definition 18 no longer suffices for a formula to define a class. In the following proposition, we also assume it not to contain class identity statements.

**Proposition 11.** *Let* $\phi$ *be elementary in the* $\psi_i$, $i \leqslant n$, *and assume that it does not contain any subformula of the form* $a = b$ *for* $a$ *or* $b$ *an abstraction term or variable. We have:*

$$\mathfrak{N}(IH_\infty) \models Cl(\hat{x}\psi_0) \wedge \ldots \wedge Cl(\hat{x}\psi_n) \rightarrow \forall y \big(y \eta \hat{z}\phi \leftrightarrow \phi(y)\big)$$

*Proof.* Recall definition 18 of elementarity (68). Let the $\psi_i$ be arbitrary, and $\phi$ elementary in them. Further assume that '=' occurs in $\phi$ only flanked by terms of the base language. Assume that $\mathfrak{N}(IH_\infty) \models$ $Cl(\psi_0) \wedge \ldots Cl(\psi_i)$.

Let o be an arbitrary object from $\omega \cup Abs$. I need to show that

$$\langle o, \hat{x}\phi \rangle \in H_\infty \Leftrightarrow \mathfrak{N}(IH_\infty) \models \phi(\overline{o})$$

Note that the left-to-right direction follows from the fact that the pair $IH_\infty$ is an admissible extension of itself (lemma 10). So it suffices to show that if $\mathfrak{N}(IH_\infty) \models \phi(\overline{o})$ then $\langle o, \hat{x}\phi \rangle \in H_\infty$.

We reason by induction on the positive complexity of $\phi$. So assume firstly that $\phi =$'$a = b$' and $\mathfrak{N}(IH_\infty) \models \phi(\overline{o})$. By our assumption about $\phi$ and without loss of generality, $a$ is the variable $x$ and $b$ is a base language term denoting in $\mathfrak{N}$ a natural number $n$. Hence, if $\mathfrak{N}(IH_\infty) \models$ $\phi(\overline{o})$ then $o = n$, and every admissible J contains $\langle o, n \rangle$. Consequently, $\langle o, \hat{x}\phi \rangle \in H_\infty$, as desired.

Secondly, let $\phi$ be an atomic formula with the relation symbol '$\eta$'. Since it is assumed to be elementary in the $\psi_i$, we know that $\phi$ is of the form $x \,\eta\, \hat{y}\psi_i$. We assume $\mathfrak{N}(IH_\infty) \models \overline{o} \,\eta\, \hat{y}\psi_i$. Hence, $\langle o, \hat{y}\psi_i \rangle \in$ $H_\infty$, such that for every admissible extension $J, K$ of $IH_\infty$, $\mathfrak{N}(J, K) \models$ $x \,\eta\, \hat{y}\psi_i$, as desired.

Still at the base of our induction on positive complexity, we now turn to negations $\phi$. By the elementarity of $\phi$, however, this implies that it is either (i) of the form '$x \not\eta\, \hat{y}\psi_i$', for some $i \leqslant n$, or (ii) of the form '$x \neq a$' (without loss of generality).

If (ii) then $\mathfrak{N}(IH_\infty) \models \phi(\overline{o})$ only if $\overline{o}$ and $a$ are both terms of the base language and we reason as just as with atomic equationhe before. So assume (i) that $\phi$ is of the form '$x \not\eta\, \hat{y}\psi_i$', and assume that $\mathfrak{N}(IH_\infty) \models \phi(\overline{o})$. Let $(J, K)$ be any admissible extension of $IH_\infty$. I need to show that $\mathfrak{N}(J, K) \models \overline{o} \not\eta\, \hat{y}\psi_i$. Since $\mathfrak{N}(IH_\infty) \models Cl(\hat{y}\psi_i)$, $\mathfrak{N}(IH_\infty) \models \forall x(x \,\eta\, \hat{y}\psi_i \vee x \,\eta\, \hat{y}\neg\psi_i)$. But we assume that $\mathfrak{N}(IH_\infty) \models$ $\phi(\overline{o})$, i.e. $\mathfrak{N}(IH_\infty) \models \overline{o} \,\eta\, \hat{y}\neg\psi_i$, hence $\langle o, \hat{x}\neg\psi_i \rangle \in H_\infty$. Therefore, $\langle o, \hat{x}\neg\psi_i \rangle \in K$, too. Since it is consistent, however, , $\langle o, \hat{y}\neg\psi_i \rangle \notin K$, hence $\mathfrak{N}(J, K) \models x \not\eta\, \hat{y}\psi_i$, as desired.

Having thus completed the base case of our induction, we proceed to the case of disjunctions $\phi$. Assume that $\mathfrak{N}(IH_\infty) \models \phi(\overline{o})$, i.e. $\mathfrak{N}(IH_\infty) \models \zeta(\overline{o})$ or $\mathfrak{N}(IH_\infty) \models \xi(\overline{o})$, for some $\zeta$,$\xi$. Assume, without loss of generality, that $\mathfrak{N}(IH_\infty) \models \zeta(\overline{o})$. Note that $\zeta$, too, is elementary in the $\psi_i$. Hence, by our induction hypothesis, $\langle o, \hat{x}\zeta \rangle \in H_\infty$. Therefore, for every admissible extension $(J, K)$ of $IH_\infty$, $\mathfrak{N}(J, K) \models \zeta(\overline{o})$. By logic, for every such $(J, K)$, $\mathfrak{N}(J, K) \models \zeta(\overline{o}) \vee \xi(\overline{o})$. Hence, $\langle o, \hat{x}\phi \rangle \in H_\infty$, as desired.

Finally, assume that $\phi =$'$\exists y\big(\zeta(x, y)\big)$' and that there is some $p$ such that $\mathfrak{N}(IH_\infty) \models \zeta(o, \overline{p})$. Then, by our induction hypothesis, $\langle o, \hat{x}\zeta(x, \overline{p}) \rangle \in H_\infty$. By reasoning just analogous to before, we conclude that $\langle o, \hat{x}\phi \rangle \in H_\infty$. □

A formula $x = \hat{y}\phi$, however, is not ensured to define a class whenever $\phi$ does. In fact, the situation is even worse. HC itself tells us that $x = \hat{x}\phi$ does not define a class whenever $\phi$ does.

**Proposition 12.** *For every formula* $\phi(\overrightarrow{x_{n+1}})$

$$\mathfrak{N}(IH_\infty) \models \forall \overrightarrow{y_n}\Big(\mathrm{Cl}(\hat{x}\phi(x, \overrightarrow{y_n})) \to \neg\mathrm{Cl}\big(\hat{y}(y = \hat{x}\phi(x, \overrightarrow{y_n}))\big)\Big)$$

Thus, the natural way of defining the singleton of a given class fails. We would both like our class theory to recover a significant fragment of naive class comprehension and its classes to be extensional. The direct approach of the present section has solved the problem of extensionality, but its theory HC violates the desideratum of comprehension badly.

**Lemma 11.** *Let* $s$ *be the* $\mathcal{L}^\wedge$*-term* $\hat{x}(x\eta x)$*. We have that neither* $\langle s, \hat{x}(x\,\eta\,x) \rangle \in H_\infty$ *nor* $\langle s, \hat{x}(x \not\eta x) \rangle \in H_\infty$*. Hence,* $x\eta x$ *does not define a class.*

*Proof of proposition 12.* Let $\phi$ be any formula and $q_0, \ldots, q_n$ any sequence of objects from the domain, with their canonical $\mathcal{L}^\wedge$-names $\overline{q_0}, \ldots, \overline{q_n}$ such that

$$\mathfrak{N}(IH_\infty) \models Cl\big(\hat{x}\phi(x, \overline{q_0}, \ldots, \overline{q_n})\big) \tag{39}$$

and let $\psi$ be the formula $x = x \wedge s\,\eta\,s$, for $s$ as in lemma 11. I show that $\langle \hat{x}\psi, \hat{y}\big(y = \hat{x}\phi(x, \overline{q_0}, \ldots, \overline{q_n})\big) \rangle \notin H_\infty$ and $\langle \hat{x}\psi, \hat{y}\big(y \neq \hat{x}\phi(x, \overline{q_0}, \ldots, \overline{q_n})\big) \rangle \notin H_\infty$, hence $\mathfrak{N}(IH_\infty) \models \neg Cl\big(\hat{y}(y = \hat{x}\phi)\big)$. I suppress the parameters $q_0, \ldots, q_n$ for the rest of the proof.

To show the first conjunct assume, for contradiction, that $\langle \hat{x}\psi, \hat{y}(y = \hat{x}\phi) \rangle \in H_\infty$. Then $\hat{x}\psi = \hat{x}\phi$ must be true in every admissible extension of $\mathfrak{N}(IH_\infty)$. In particular, the pair $\langle \hat{x}\psi, \hat{x}\phi \rangle$ must be in the fixed point identity relation $I_\infty$ (cf lemma 10). By its fixed point character, we have

$$\forall J \forall K\Big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \forall x\big(x = x \wedge s\eta s \leftrightarrow \phi(x)\big)\Big) \tag{40}$$

Now let $o$ be any object in the domain of $\mathfrak{N}(IH_\infty)$; that is, let $o$ be a number or an abstraction term. Since $\mathfrak{N}(IH_\infty) \models \forall y(y\eta\hat{x}\phi \vee y\eta\hat{x}\neg\phi)$, we can assume that either $\langle o, \hat{x}\phi \rangle \in H_\infty$ or $\langle o, \hat{x}\neg\phi \rangle \in H_\infty$. Firstly, assume that $\langle o, \hat{x}\phi \rangle \in H_\infty$. Hence $\forall J \forall K\big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models \phi(\overline{o})\big)$. Then by (40) and logic, $\forall J \forall K(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J, K) \models s\eta s)$ Hence, $\langle s, s \rangle$ must be in $H_\infty$, contrary to lemma 11.

Secondly, assume that $\langle o, \hat{x}\neg\phi\rangle \in H_\infty$ such that $\forall J \forall K(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \neg\phi(\bar{o}))$. Then by (40) and logic, $\forall J \forall K(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \neg\, s\, \eta\, s)$. This requires, again contrary to lemma 11, that $\langle s, \hat{x}(x \not\eta x)\rangle \in H_\infty$.

To show the second conjunct assume, for contradiction, that $\langle \hat{x}\psi, \hat{y}(y \neq \hat{x}\phi)\rangle \in H_\infty$. By the fixed point character of the pair $IH_\infty$, $\forall J \forall K(IH_\infty \Subset J, K \Rightarrow \langle \hat{x}\psi, \hat{x}\phi\rangle \notin J)$ This is the case only if either (i) $\langle \hat{x}\psi, \hat{x}\neg\phi\rangle \in I_\infty$ or (ii) $\langle \hat{x}\neg\psi, \hat{x}\phi\rangle \in I_\infty$.

Assume (i), such that

$$\forall J \forall K\Big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \forall x\big(x = x \wedge s\eta s \leftrightarrow \neg\phi(x)\big)\Big) \quad (41)$$

Let o be any object of the domain. As before, since $\mathfrak{N}(IH_\infty) \vDash Cl(\hat{x}\phi)$, we have that either $\langle o, \hat{x}\phi\rangle \in H_\infty$ or $\langle o, \hat{x}\neg\phi\rangle \in H_\infty$. Firstly, assume the former. Then

$$\forall J \forall K\big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \phi(\bar{o})\big) \quad (42)$$

Lemma 11 ensures that there is a pair $(J_0, K_0) \sqsupseteq IH_\infty$ such that $\langle s, s\rangle \in K_0$. By (41), $\mathfrak{N}(J_0, K_0) \vDash \bar{o} = \bar{o} \wedge s\eta s \leftrightarrow \neg\phi(\bar{o})$ and by (42) and logic, $\mathfrak{N}(J_0, K_0) \vDash \neg\, s\eta s$ which contradicts our assumption that $K_0$ contains $\langle s, s\rangle$.

Secondly, assume that $\langle o, \hat{x}\neg\phi\rangle \in H_\infty$. Now choose $(J_1, K_1) \sqsupseteq IH_\infty$ such that $K_1$ does not contain $\langle s, s\rangle$ ($IH_\infty$ itself is such a pair). By (41), $\mathfrak{N}(J_1, K_1) \vDash x = x \wedge s\eta s \leftrightarrow \neg\phi(\bar{o})$. By our assumption and logic $\mathfrak{N}(J_1, K_1) \vDash s\eta s$ contrary to our choice of $K_1$.

Now assume (ii), such that

$$\forall J \forall K\Big(IH_\infty \Subset J, K \Rightarrow \mathfrak{N}(J,K) \vDash \forall x\big(\neg(x = x \wedge s\eta s) \leftrightarrow \phi(x)\big)\Big) \quad (43)$$

We reason just conversely. For any o we firstly assume $\langle o, \hat{x}\phi\rangle \in H_\infty$ and choose a $J_0$ containing $\langle s, s\rangle$. (43) implies that $\mathfrak{N}(J_0, K_0) \vDash \neg\, s\eta s$, contradiction. Secondly, we assume $\langle o, \hat{x}\neg\phi\rangle \in H_\infty$, choose $J_1$ *not* containing $\langle s, s\rangle$, which contradicts our assumption and (43).

It remains to show lemma 11. Assume otherwise, then either (†) $\langle \hat{x}(x\eta x), \hat{x}(x\eta x)\rangle \in H_\infty$ or (‡) $\langle \hat{x}(x\eta x), \hat{x}(x \not\eta x)\rangle \in H_\infty$. If (†) then there's a least ordinal $\alpha + 1$ such that $\langle \hat{x}(x\eta x), \hat{x}(x\eta x)\rangle \in H_{\alpha+1}$. That is, for all $J, K$ admissibly extending $H_\alpha$, the sentence $\hat{x}(x\eta x)\eta\hat{x}(x\eta x)$ is true in the model $\mathfrak{N}(J,K)$. In particular, it must be true in $\mathfrak{N}(IH_{\alpha+1})$, hence $\langle \hat{x}(x\eta x), \hat{x}(x\eta x)\rangle$ is found already in $H_\alpha$, contradiction.

If (‡), then there's a least ordinal $\alpha + 1$ such that $\langle \hat{x}(x\eta x), \hat{x}(x \not\eta x)\rangle \in H_{\alpha+1}$. So, for every admissible extension $J, K$ of $H_\alpha$, $\langle \hat{x}(x\eta x), \hat{x}(x\eta x)\rangle$ is not in J. This, however, can only be if $\langle \hat{x}(x\eta x), \hat{x}(x \not\eta x)\rangle \in H_\alpha$, since otherwise the equivalent closure of $H_\alpha \cup \{\langle \hat{x}(x\eta x), \hat{x}(x\eta x)\rangle\}$ is an admissible, in particular consistent, extension, contrary to what we have just said. If this is so, however, $\alpha + 1$ is not the least ordinal containing $\langle \hat{x}(x\eta x), \hat{x}(x \not\eta x)\rangle$, contradiction. $\square$

## 4.6 CONCLUSION

In this chapter, I examined the prospects of class theory inspired by theories of grounded truth. I asked how to restrict the schema of class comprehension to grounded formulae, just as Kripke restricted Tarski's T-schema to grounded sentences.

Having laid out desiderata, I first explored the derivative approach. I translated "$x$ is in the class of the $\phi$s" as "$\phi(x)$ is true" (p. 66). Through this translation, a theory of grounded truth induces a corresponding theory of grounded classes. The desiderata of section 4.2 suggested to start from the theory of truth KFB. The resulting class theory HKFB is closed under classical logic, allows for arithmetical as well as set-theoretical base theories, and proves comprehension for every *elementary* formula (p. 71). However, HKFB does not satisfy the desideratum of extensionality (§ 4.4).

In section 4.5 I turned to developing a theory of grounded classes directly. I described the extension of an arbitrary base model by a relation of grounded class membership and a relation of grounded class identity. The resulting model provides a theory HC whose classes are extensional in the strict sense that firstly, HC identifies $\hat{x}\phi$ and $\hat{x}\psi$ just in case $\forall y(y \,\eta\, \hat{x}\phi \leftrightarrow y \,\eta\, \hat{x}\psi)$. Secondly, classes that HC identifies are indiscernible in the theory. However, these positive results are blighted by a severe deficiency: according to the theory HC, whenever $\phi$ defines a class, $x = \hat{y}\phi$ does not.

Prima facie, we would like classes to be extensional, and their theory to provide natural ways of defining classes. My findings cast doubt on whether both can be achieved by the groundedness approach to class theory.

# 5

# WHAT IS THE PHILOSOPHICAL SIGNIFICANCE OF GROUNDEDNESS?

## 5.1 INTRODUCTION

In chapter 2 I presented a general, formal concept of groundedness. Then, I discussed applications of this general concept: the iterative conception of sets (§2.7), grounded truth (§3) as well as approaches to a grounded theory of classes (chapter 4). Now, I take a step back and ask for the philosophical significance of this formal concept of groundedness.

This question is not easy to answer. I will argue that many instances of the general concept lack philosophical content. For others, it is at least controversial whether they have such. I will give examples in sections 5.2 to 5.4 below. Together, I take them to be evidence that the general, formal concept of groundedness from chapter 2 is in need of philosophical supplementation. In the next chapter, I will then present one way of accounting for the significance of one specific case of groundedness.

What is philosophical significance, and philosophical content? I do not use these expressions in any deeper sense than for what philosophers tend to agree on as worth their attention. This certainly is not a precise notion. However, I believe that every trained practitioner of our discipline has sufficient grasp of it. Presumably, our access is through examples. So let me give some. The biological concept of evolution has gained philosophical significance, witness for example Sober [1984], while the same cannot be said of, say, metabolism.

Or, Frege's theorem, that in second-order logic, Peano Arithmetic can be derived from Hume's Principle (see, e.g., Heck [1993]), is philosophically significant while Gonthier's 2005 result, that the four colour theorem can be proved in the type theory proof assistent *Coq* [Gonthier [2008]], is not. As a further example, more closely to the present topic, the semantic paradoxes, Epimenides, Curry and others, are philosophically significant while 21-year-old Frederic having had only five birthdays, arguably is not [Quine, 1976, p. 1].

Returning to groundedness, I ask: is it philosophically significant?

## 5.2 FORSTER'S ITERATIVE CONCEPTION OF CHURCH-OSWALD CLASSES

I motivated my concept of groundedness as further generalization of Forster's 2008 take on the iterative conception . However, the philosophical significance of his main example is controversial. It is the Church-Oswald construction of models for theories with a universal class [Forster, 2008, §§2,5]. Their classes can be viewed as grounded in the sense of my formal definition, but it is not obvious whether this case of groundedness is philosophically significant. I briefly rehearse the simplest Church-Oswald construction in the usual, set-theoretic setting before I explain how it exemplifies groundedness.[1] Then, I will argue that the philosophical significance of this instance of groundedness is contentious.

Take any model $\langle M, E\rangle$ of Zermelo-Fraenkel set theory ZF and attach labels, say 0 and 1, to the objects of its domain. For example, this is implemented by taking pairs $\langle x, 0\rangle$, $\langle x, 1\rangle$ for $x \in M$. Choose a bijection $c$ that maps every object in $M$ to exactly one such pair, not necessarily containing this set itself. That is, $c(x)$ is a pair $\langle y, 0\rangle$ or $\langle y, 1\rangle$ for some $y \in M$. We assume that the rank of $c(x)$ is greater than that of $x$.

The function $c$ allows us to define a relation $F$ on $M$. Together with the domain $M$ of the original model, this new relation gives rise to a model $\langle M, F\rangle$ of a theory of classes with a universal class.

**Definition 23.** For $x, y \in M$

$$xFy :\Leftrightarrow \exists z \begin{cases} c(y) = \langle z, 0\rangle \text{ and } x \in z \\ \text{or} \\ c(y) = \langle z, 1\rangle \text{ and } x \notin z \end{cases}$$

Let $\mathcal{L}$ be a basic language of class theory (recall the previous chapter), the language of first-order logic extended by the relation symbol '$\eta$'. $\langle M, F\rangle$ is an $\mathcal{L}$-structure. $F$ functions as a relation of class membership, and the object of the domain $M$ function as classes. In particular, the object $u \in M$ such that $c(u) = \langle \varnothing, 1\rangle$ functions as a universal class: every $x \in M$ bears $F$ to $u$. It can be shown that the model $\langle M, F\rangle$ validates extensionality with respect to the membership relation $F$.

In his 2008 article, Forster provides an alternative characterization of these classes (§2). It bases on two generators.[2] The first one is well known. It is simply the set-generator **S** from section 2.7 (definition 12). The other is rather unusual: it allows us to generate from some things $xx$ the class of everything that is *not among* $xx$. This generator is defined as follows.

**Definition 24** (Forster's Complement Class Generator)**.** $y$ is $\mathbf{Ƨ}$ generated from $yy$ iff has as its members all and only the $z$ which are not one of $xx$.

The classes of the model $\langle M, F\rangle$ can be viewed as generated from their elements (in the sense of the relation $F$), by **S**, or from those objects that are *not* their elements, by $\mathbf{Ƨ}$. To see this, recall that for every $y \in M$, there is a $z \in M$ such that either $c(y) = \langle z, 0\rangle$ or $c(y) = \langle z, 1\rangle$. In the first case, the model says that $x$ is an element of $y$, for every $x$, if and only if $x \in z$. That is, the theory of $\langle M, F\rangle$ takes $x$ to be

1 My exposition follows closely Forster's [Forster, 2008, §5], but see also Oswald [1976].

2 Ingeniously, Forster speaks of 'wands'.

an element of $y$ if and only if $x$ is among those things that we, in the meta-theory, know to be an element of $z$. In other words, $y$ is the class of the things in $z$; $z$ collects the objects $zz$ from which $y$ is generated by **S**.

In the second case when $c(y) = \langle z, 1 \rangle$, the model says that $x$ is an element of $y$ if and only if $x \notin z$. In other words, the object theory takes $x$ to be an element of $y$ if and only if $x$ is *not* among the things that the meta-theory knows to be in $z$. This time, therefore, $z$ represents the objects $zz$ from which $y$ is generated by **Ƨ**. Note that unlike a standard set, a Forster class may be not grounded in its members, but in those things which are precisely not its members, if **Ƨ** generated from them.

Consequently, the sets of the model $\langle M, F \rangle$ represent objects of a new kind, classes that are generated in a way quite unlike the standard generator of sets. I will use the label 'Forster class' to refer to these objects stipulated by Forster's new interpretation of the Church-Oswald models.

Forster argues that the **Ƨ**-**S**-grounded classes are as legitimate as the standard, **S**-grounded sets. The fact that he considers it necessary to add philosophical argument to his characterization of the Forster classes as **Ƨ**-**S**-grounded, already suggests that the philosophical content of this characterization is not obvious. In the remainder of this section I will discuss whether Forster succeeds in establishing that **Ƨ**-**S**-groundedness is as legitimate as **S**-groundedness. I start with a series of indirect arguments that Forster gives, as they will help clarifying what is at stake.

Forster discusses three worries one may have about the **Ƨ**-**S**-groundedness characterization of the Forster classes [Forster, 2008, §4]. These worries are not of mathematical nature. It is not questioned that there are Church-Oswald models, nor that the classes of these models are **Ƨ**-**S**-grounded. Instead, is doubted that Forster's two-generator picture is as philosophically significant as the standard, one-generator picture of the cumulative hierarchy [Forster, 2008, "Horn 1" on p. 108]. Thus, the objections Forster considers are worries about the philosophical significance of **Ƨ**-**S**-groundedness.

Forster phrases these objections as arguments that the **Ƨ**-**S**-grounded objects are not sets. I do not think that this is the most felicitous way of putting it. After all, it is trivial that among the **Ƨ**-**S**-grounded things there are objects that are not sets (in the standard sense). As we have observed above, there is an object $u \in M$ which our simple Church-Oswald model $\langle M, F \rangle$ treats as a universal class, and there is no universal set. Unless, of course, by 'set' we no longer mean the **S**-grounded objects of the standard cumulative hierarchy, but allow for a more liberal use of this expression; in particular, unless we start calling the Forster classes 'sets'. However, this is not what is disagreed on. Forster does not engage in a merely verbal dispute. Therefore, I understand the objections considered by Forster, as arguments that whereas **S**-groundedness provides a philosophical case for sets, **Ƨ**-**S**-groundedness does not do the same thing for Forster classes. Does he succeed in fending off these objections?

The first argument goes as follows [Forster, 2008, §4.1]. **S**-grounded sets are constituted from their elements, but **Ƨ**-**S**-grounded classes are not. Therefore, **S**-groundedness is significant, while the **Ƨ**-**S**-groundedness is not.

Forster responds to this argument in two steps. Firstly, he argues that to say that sets are constituted from their elements is just to say that sets are extensional. Secondly, he points out that the Forster classes are extensional, too. Therefore, **S**-grounded sets and **Ƨ**-**S**-grounded classes do not differ after all in the relevant sense.

I do not think that Forster's response is conclusive. There is a relevant sense in which sets are constituted from their elements, a sense which is not exhausted by the extensionality of sets. In his seminal 1971 article, Boolos explicitly contrasts two characteristics of sets: on the one hand, their extensionality, on the other hand, the fact that '[...] the elements of a set are "prior" to it' (p. 216). To say that a set is constituted from its elements may mean that it has both of the characteristics mentioned by Boolos, only that its elements are prior to it, or finally just that it is extensional. Forster's response addresses this latter sense in which a set is constituted from its elements, but not the others. The argument that **Ƨ**-**S**-groundedness is insignificant, however, can equally be formulated based on the other two senses. In particular, it is plausible to say that the standard set generator **S** tracks the priority of some things to their set, while Forster's complement generator **Ƨ** does not. Further, it can well be argued that the philosophical significance of **S**-groundedness stems from the fact that a set is **S**-grounded in precisely the things that are prior to it [Potter, 2004, §3.3]. The next chapter (§ 6.2) will pick up this line of thought and develop it further.

The second argument Forster considers is based on the following observation. Given some things *zz*, the condition of not being among *zz* only defines a plurality if the universe is already given as a definite collection. Otherwise, it is not definite which members a Forster class has.[3]

However, it is contentious to assume that the universe is a definite plurality.[4] In fact, it conflicts with our assumption about the set generator **S**. To see how, let $uu$ be all the things there are and use **S** to generate from $uu$ the set of all things $\{uu\}$, contradiction. Therefore, *prima facie* what members an **Ƨ**-generated class has is not a definite matter. In this sense, a Forster class is an *intension*, not an *extension*.

3 Note that I specify slightly Forster's own exposition, in that I explicate his temporal metaphor of "the end of time" in terms of whether or not the universe is definite. Concerning the relevant concept of definiteness, recall p. 38.

4 See the extensive literature on *absolute generality*, e.g. in Rayo and Uzquiano [2006].

Standard, **S**-grounded sets are extensions. The elements of a given set are just those things from which it is generated, and therefore always ensured to be definite – otherwise, the set could simply not have been generated. Therefore, whereas **S**-groundedness ensures having a definite range of elements, **Ƨ**-**S**-groundedness does not. Hence, the argument goes, the philosophical significance of **S**-grounded sets does not carry over to the **Ƨ**-**S**-grounded Forster classes.

In Forster's discussion of this objection [§4.2] I discern two distinct, indeed possibly conflicting, responses. On the one hand, Forster accepts that the fact that unlike standard sets, a Forster class generated through **Ƨ** has definite members only if the universe is definite, marks '[...] an important difference' between **S**-groundedness and **Ƨ**-**S**-groundedness [Forster, 2008, p. 105]. It is not a mathematical difference, since the relevant notion of definiteness is of philosophical nature. Hence, Forster acknowledges at least one aspect in which his generalized iterative conception does not ensure philosophical significance.

On the other hand, Forster argues that the objection overshoots. It does not only cast doubt on the legitimacy of **Ƨ**-**S**-groundedness, but on the legitimacy of inductive, or in Forster's terms, recursive constructions quite generally. Thus, the objection contradicts what Forster labels *Conway's principle*, that 'objects may be created from earlier objects in any reasonably constructive fashion' [Forster, 2008, p. 99].[5]

Why should Forster's opponent be moved by Conway's principle? It depends on what we take it to mean. If the principle says that every inductive definition ensures philosophical significance, then for Forster to uphold it, is not to provide an argument for, but simply to repeat his conviction that the **Ƨ**-**S**-groundedness of Forster classes is as significant as standard **S**-groundedness.

A more charitable reading of Conway's principle is as giving expression to a feature of mathematical reasoning, namely that some collection of things having been defined inductively licences reference to them. On this reading, Forster's response becomes that disallowing Forster classes contradicts mathematical practice. This would certainly be unacceptable.

However, the objection is not that **Ƨ**-**S**-groundedness does not licence mathematical reasoning with Forster classes. Their mathematical significance is already accounted for by the standard Church-Oswald model construction (definition 23). Instead, the objection is that Forster classes do not have the same philosophical significance as standard sets. Therefore, the objection does not contradict Conway's principle as suggested by Forster.

In sum, Forster's reference to Conway's principle either merely restates his view that **Ƨ**-**S**-groundedness is philosophically as good as

5 Forster cites Conway [2001].

standard **S**-groundedness, or it reminds us of the fact that in mathematical reasoning, inductive definition licences reference. The former is not an argument, while the latter does not conflict with denying its philosophical significance. Therefore, Forster's argument that the objection from definiteness overshoots, is not conclusive.

The third objection that Forster considers is a slippery slope argument. It goes as follows. If we accept that Forster's conception of classes, based on the set generator **S** as well as the complement generator **Ƨ**, is as significant as the received iterative conception of sets, based on **S** alone, then any other generator has equal claim to produce legitimate objects.

One philosopher's modus tollens is the other's modus ponens. Forster is ready to accept that for any generator $\gimel$, $\gimel$-groundedness is philosophically significant. More importantly, however, he points out that even if we did not accept every generator, the threat of regress [Forster, 2008, p. 106]

> [...] is not *by itself* an argument for drawing the line so close to home that [...][**Ƨ**-**S**-groundedness] is excluded.

I agree. By a similar thought, however, the fact that we cannot as suggested argue against **Ƨ**-**S**-groundedness is no reason to agree with Forster. The burden of proof is on him to show that his conception of Forster classes is philosophically as good as the standard iterative conception of sets. After all, it is the received view that standard set theory Z, possibly ZF, receives good motivation from the iterative conception, and that in this respect it is superior to alternative theories, such as that of the Church-Oswald models. Forster claims that the Church-Oswald theory is just as well motivated. However, unless he provides positive reason for this, methodology requires us to adhere to the received view.

So far, I have only presented Forster's indirect arguments, by which he responds to objections likely to be put forward against his unconventional view. In fact, however, Forster also provides a positive argument. Indeed, the first two sections of his paper are well viewed as arguing that his two-generator iterative conception of Forster classes provides them with as good philosophical motivation as does the one-generator iterative conception, i.e. **S**-groundedness, for standard set theory.

Forster's argument rests on the following assumption [Forster, 2008, p. 98].

(Q) The appeal of the cumulative hierarchy lies precisely in its neat response to Quine's challenge.

By 'the cumulative hierarchy' Forster refers to what I call **S**-groundedness. The 'appeal' that Forster ascribes to it is its appeal to philosophers, hence, at least partly, its philosophical significance.

By ‘Quine’s challenge’, Forster means Quine’s famous insight that a theorist can use first-order logic with identity to reason about things of a certain kind, only if she has an identity criterion for them. The cumulative hierarchy, or rather the view that every set is found at some stage of it (i.e. is **S**-grounded), satisfies this necessary condition and responds to Quine’s challenge, because it provides an identity criterion for sets. Finally, this response is ‘neat’ in the precise sense that the identity criterion provided comes with a ‘[. . . ] recursive algorithm for deciding identity’ [Forster, 2008, p. 98].

In sum, Forster’s assumption (Q) is well paraphrased as follows.

(Q′) The philosophical significance of **S**-groundedness is that it gives a recursive identity criterion for sets.

Forster points out that **Ƨ**-**S**-groundedness, too, provides a recursive identity criterion for Forster classes. Based on the assumption (Q′), he concludes that **Ƨ**-**S**-groundedness is philosophically as significant as **S**-groundedness. I do not accept this conclusion, because I reject Forster’s premise (Q′). In the remainder of this section, I will argue that (Q′) is false, and conclude that Forster’s positive argument for the significance of **Ƨ**-**S**-groundedness does not go through.

At least on one reading, (Q′) presupposes that any recursive identity criterion renders its domain philosophically significant. This, however, is not true. I give a case in which we have an algorithm to establish identity and distinctness facts in some given domain, but which lacks philosophical significance. Imagine a rather exclusive international spy ring. It consists of two individuals, Mary and Harry. As all spies they have their cover identities. Mary goes under the names ‘Alina’ and ‘Varvara’, while Harry is known both as ‘Valentin’ and as ‘Nikita’. Now, under each of these covers, Mary and Harry infiltrate the enemy’s secret service. It, too, assigns them cover idenities, not knowing that it is dealing with double agents. For example, the service assigns its agent Varvara, who really is Mary, one cover identity under the name ‘Sarah’, and another under the name ‘Laura’. As a result, each of our spies is endowed with two layers of cover identites, as depicted in figure 13.

We may imagine Laura, Sarah and the others again to be assigned double agent missions, and thus be given further, more remote cover names. But I can make my point without them. There is an algorithm for Mary’s and Harry’s superiors to identify who of their newly hired agents is in fact which a *triple*-agent of their own. This procedure is given by the following recursion. Mary is Mary, and Harry is Harry; and to answer whether $x$ is $y$, uncover $x$’s identity $u$ and $y$’s identity $v$, and ask if $u = v$.

For example, it may be asked if Janet and Sarah are the same triple agent. Well, following our algorithm, we firstly ask if Alina is Varvara. In order to answer this question, we again apply the step of our recur-

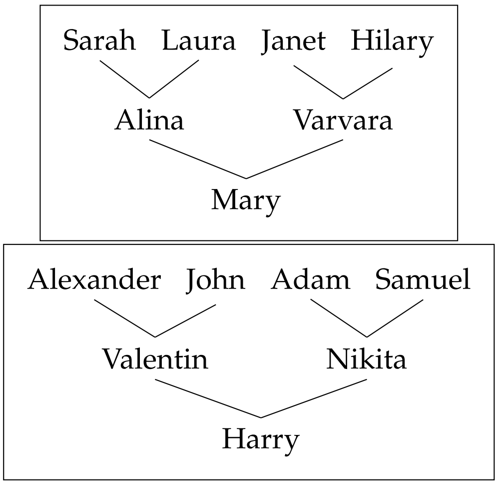


Figure 13: Mary's and Harry's cover-identities

sion, and find that Alina and Varvara both are Mary. By the recursion base, however, Mary is Mary, and we conclude that Janet is Sarah.

Thus, we have a recursive identity criterion for this group of international triple agents. However, they are not philosophically significant. I conclude that a recursive identity criterion by itself does not account for philosophical significance, contrary to Forster's (Q′). Consequently, the significance of **S**-groundedness still requires explanation.

It may be objected that this last step of my argument fails because (Q′) does not imply that recursivity of identity is sufficient for philosophical significance. I assumed that (Q′) implies that much, but there may be a more charitable reading of it which does not have such strong implications, but accounts for the significance of **S**-groundedness nonetheless. Therefore, the objection goes, I have failed to fend off Forster's case.

In response to this challenge, I admit that I have not addressed every reading of (Q′). Doing so would require spelling out what precisely it means for the significance of **S**-groundedness to be that **S**-groundedness 'gives a recursive identity criterion'. All I have shown is that it cannot mean that the latter suffices for significance. However, no clear alternative reading is readily available, in terms of which Forster's case for the significance of groundedness can be recovered. Even if I have not given a deductive argument, therefore, I have at least cast doubt on (Q′), and shifted the burden on Forster to explain just how the significance of **S**-groundedness consists in it giving a recursive identity criterion. Forster's positive argument for the significance of **Ƨ**-**S**-groundedness is that it provides us with a recursive identity criterion for Forster classes. Thus, I have also given reason to doubt the philosophical significance of Forster's **Ƨ**-**S**-groundedness, and of the formal concept of groundedness in general.

## 5.3 FRIEDMAN AND SHEARD'S MODELS OF TRUTH

In the previous section I have argued that the philosophical significance of **Ƨ-S**-groundedness is controversial. Thus, I have found reason to be skeptical about the philosophical significance of the general, formal concept of groundedness from chapter 2 In this section, I will give another case of groundedness whose philosophical significance is not obvious. In fact, I now turn to a case that, unlike Forster's conception of classes, was never intended as philosophically significant.

Recall, from chapter 3, Kripke's notion of *grounded truth*. It comes in several variants, each based on a distinct monotone evaluation scheme. The Strong Kleene variant (§3.5) has received the most attention, but semantic groundedness based on Weak Kleene, or on a supervaluational scheme have also been discussed.

I would like to emphasize that these cases of groundedness are philosophically significant. They have been discussed in philosophical journals and books, and not merely so in the wake of Kripke's seminal paper, but repeatedly over the past four decades. Today, Kripke's theory of truth, or family of theories to be precise, has become the standard type-free theory of truth, to the extent that such consensus is found among philosophers. In particular, it is considered to have advantages over its revision-theoretic contenders. The appeal that Kripke's theory has to the majority of philosophers is at least partly due to that it is motivated from his notion of groundedness, which is an instance of the general formal concept from chapter 2. Therefore, semantic groundedness, in its Strong or Weak Kleene variant, or in one of its supervaluational variants, is philosophically significant.

However, semantic groundedness has further variants of which this cannot be said. They are found in another seminal piece of formal theory of truth, Friedman and Sheard [1987]. Friedman and Sheard provide an impressive array of results as to which axioms and rules, each of which embodies some aspect of naïve truth, are mutually consistent. For this purpose, they construct models very similar to Kripke's. However, these models themselves are not intended to capture an aspect of truth. They are merely technical devices to show that certain axioms are consistent. Nevertheless, their truth is grounded much like Kripke's (see p. 44).

For example, Friedman and Sheard construct a model $\mathfrak{N}(\mathrm{Th}_\infty)$ whose truth predicate $\mathrm{Th}_\infty$ is the union of a sequence of sets $\mathrm{Th}_n$, where $\mathrm{Th}_0$ is true first-order arithmetic, and $\mathrm{Th}_{n+1}$ is the set of sentences $\phi$ such that [Friedman and Sheard, 1987, §3, G]

$$\Lambda \cup \{T\ulcorner\psi\urcorner : \psi \in \mathrm{Th}_n\} \cup \{T\ulcorner\forall x\psi(x)\urcorner : \forall x T\ulcorner\psi(\dot{x})\urcorner \in \mathrm{Th}_n\} \vDash_\omega \phi \quad (44)$$

Here, $\vDash_\omega$ is consequence in $\omega$-logic: classical logic, in the language of arithmetic, enhanced by the following rule.

$$\frac{\phi(\overline{0}) \qquad \phi(\overline{1}) \qquad \ldots}{\forall x \phi(x)}$$

This model validates the following axiom system whose consistency is thereby proved. [6]

$$\text{T-Intro}\ \frac{\phi}{T\ulcorner\phi\urcorner} \qquad \text{T-Elim}\ \frac{T\ulcorner\phi\urcorner}{\phi} \qquad \frac{\neg T\ulcorner\phi\urcorner}{\neg\phi}\ \neg\text{T-Elim}$$
$$\text{T-Rep}\ T\ulcorner T\ulcorner\phi\urcorner\urcorner \rightarrow T\ulcorner\phi\urcorner$$
$$\text{U-Inf}\ \forall x T\ulcorner\phi(\dot{x})\urcorner \rightarrow T\ulcorner\forall x\phi(x)\urcorner$$

The sentences in $Th_\infty$, now, can be shown to be *grounded* in the truths of arithmetic N, in a manner very similar to how the sentences of Kripke's fixed point are grounded in them. More precisely, there is a generator $\daleth$, not unlike the generators underlying Kripkean groundedness, such that $\psi \in Th_\infty$ iff $\phi$ is $\daleth$-grounded in N.

On the *high-resolution* characterization of semantic groundedness that I have found advantageous (see §3.4 above), the sentences in the least fixed point of, say, the Strong Kleene Kripke jump, are viewed as grounded in the truths of arithmetic, through the combination of the Kripke truth generator **T** (p. 49) and the Tarski logic generator **W** (p. 28). Similarly, the sentences of Friedman and Sheard's theory $Th_\infty$ can be viewed as generated from the truths of arithmetic by two generators.

The definition of $Th_{n+1}$ above (equation 44) has two key components. On the one hand, the relation $\vDash_\omega$ of consequence in $\omega$-logic; on the other hand, the step from $\forall x T\ulcorner\psi(\dot{x})\urcorner$ to $T\ulcorner\forall x\psi(x)\urcorner$, and the step from $\psi$ to $T\ulcorner\psi\urcorner$. Accordingly, we can view $Th_\infty$, too, as grounded through the combination of a logic- and a truth-generator.

The Friedman Sheard truth generator ('**F**'), on the one hand, is given by two rules. The first rule is T-Intro, in terms of which we have also characterized Kripke's truth generator **T** (p. 49). The second rule allows us to infer that it is true that for everything it is the case that $\phi$, from the assumption that for everything it is true that $\phi$.

$$\frac{\forall x T\ulcorner\phi(\dot{x})\urcorner}{T\ulcorner\forall x\phi(x)\urcorner}$$

Note that **F**, just like Kripke's truth generator **T**, is deterministic in the sense of definition 6.

6 The construction above is simpler than what Friedman and Sheard do strictly speaking. At every stage, they do not only add every $\omega$-logic consequence of $N \cup \{T\ulcorner\psi\urcorner : \psi \in Th_n\}$, but also every instance of the schemata T-Rep and U-Inf. However, for their purpose, i.e. to prove the consistency of the axiom system given above, the simpler construction discussed suffices.
Also, I suppress the fact that the base theory of $Th_\infty$ is not just first-order arithmetic PA, but also includes basic truth-theoretic principles [Friedman and Sheard, 1987, p. 4]. Such details do not matter for the general point I intend to make, that the construction exemplifies the general concept of groundedness from chapter 2.

The logic generator, on the other hand, is simply **W**; recall that it allows us to generate universal quantifications from an infinity of sentences (p. 29).

**Proposition 13.** *The sentences in* $\mathrm{Th}_\infty$ *are* ***F-W****-grounded in the truths of first-order arithmetic* $\mathbb{N}$.

$$\textit{If } \phi \in \mathrm{Th}_\infty \textit{ then } \phi \ \textbf{\textit{F-W}}\textit{-grounded in } \mathbb{N}$$

*Proof.* Since $\mathrm{Th}_\infty = \bigcup_{n<\omega} \mathrm{Th}_n$, it is natural to reason by induction on $n$. $\mathrm{Th}_0 = \mathbb{N}$, hence $\phi \in \mathrm{Th}_0$ is trivially **F-W**-grounded in $\mathbb{N}$.

For $\phi \in \mathrm{Th}_{n+1}$, we know that $\phi \in \mathbb{N}$, or $\phi$ is of the form $T\ulcorner\psi\urcorner$ and $\psi \in \mathrm{Th}_n$, or $\phi =$'$T\ulcorner\forall x\psi(x)\urcorner$' and $\forall x T\ulcorner\psi(x)\urcorner \in \mathrm{Th}_n$ or, finally, $\phi$ follows from some $\Gamma \subseteq \mathrm{Th}_n$ in $\omega$-logic. We reason by cases. Firstly, $\phi \in \mathbb{N}$ iff $\phi$ is trivially **F-W**-grounded in $\mathbb{N}$. Secondly, $\phi =$'$T\ulcorner\psi\urcorner$' for $\psi \in \mathrm{Th}_n$, iff $\psi, \mathbf{F}\phi$, which in turn by our induction hypothesis holds just in case $\phi$ is **F-W**-grounded in $\mathrm{Th}_n$. Analogously for $\phi =$'$T\ulcorner\forall x\psi(x)\urcorner$' and $\forall x T\ulcorner\psi(x)\urcorner \in \mathrm{Th}_n$. Finally, $\phi$ is an $\omega$-consequence of some $\Gamma \subseteq \mathrm{Th}_n$ just in case $\phi$ is **W**-grounded in $\mathrm{Th}_n$, hence **F-W**-grounded in $\mathrm{Th}_n$. □

So, the truth predicate of Friedman and Sheard's model $\mathfrak{N}(\mathrm{Th}_\infty)$ satisfies the general concept of groundedness. Formally, $\mathrm{Th}_\infty$ is as much a predicate of *grounded* truth as is the least fixed point of Kripke's Strong Kleene jump (section 3.5). However, its groundedness is not philosophically significant. It is not intended to be so. As to their paper, Friedman and Sheard are explicit that its approach is primarily logical, and that they do not intend to make a philosophical point [Friedman and Sheard, 1987, p. 2].

> We are not solving a problem in philosophy, but rather a problem in logic with a philosophical motivation.

As to the model constructions in section 3 of the paper, their sole purpose is to prove consistent certain collections of axioms and rules governing 'T'. No further role is mentioned nor any aspect of these constructions is discussed.

Moreover, even if we went beyond how Friedman and Sheard use their models, and sought to take them seriously as philosophers, this would still not render significant the groundedness of the sentences in $\mathrm{Th}_\infty$. Firstly, when I presented the model $\mathfrak{N}(\mathrm{Th}_\infty)$ above (44), I defined the set of sentences $\mathrm{Th}_\infty$ in a manner that renders it easy to see their groundedness, starting from $\mathbb{N}$ and step by step adding sentences with 'T'. Friedman and Sheard, however, define it explicitly as the least set containing those axioms and closed under those rules whose consistency they want to prove. Only in passing they remark that $\mathrm{Th}_\infty$ can also be defined as I did above. Therefore, even if Friedman and Sheard's construction of the model $\mathfrak{N}(\mathrm{Th}_\infty)$ had philosophical significance, it would not obviously carry over to its groundedness.

Secondly, $\mathfrak{N}(\mathrm{Th}_\infty)$ is merely one of a list of models each of which validates a specific axiomatic system. We have as little reason to believe in the philosophical significance of $\mathfrak{N}(\mathrm{Th}_\infty)$ as in the relevance of any of the others. However, many of these other models do not exhibit groundedness. For example, Friedman and Sheard use, under the heading of "converging" truth, revision-theoretic tools to construct a model that validates the inference from $\phi$ to $T\ulcorner\phi\urcorner$ and vice versa [Friedman and Sheard, 1987, §3.D].[7] Its truth predicate $\mathrm{Th}'_\infty$ cannot be read as a predicate of grounded truth in the same way as I have found $\mathrm{Th}_\infty$ to be grounded. Therefore, even if we had reason to believe in the philosophical significance of $\mathrm{Th}_\infty$, it would not automatically be reason to take its *groundedness* to be significant.

I conclude that Friedman and Sheard's model $\mathfrak{N}(\ulcorner\mathrm{Th}_\infty\urcorner)$ is a case of groundedness that is not intended to be philosophically significant, that there is no reason to assume it is, and that the attempt of arguing for its significance faces difficulties. Thus, I have given additional evidence that the general, formal concept of groundedness from chapter 2 is in need of philosophical supplementation.

## 5.4 HOW TO GROUND ANYTHING

In the previous two sections I have presented cases of groundedness each of which resembles a paradigmatic, and philosophically significant, instance of the general concept (sets respectively truth) but whose philosophical significance is at least contentious. I now turn to present cases that satisfy the general theory, but do not even resemble anything philosophically significant. I show how to, speaking informally, *cook up* groundedness, and thus produce many cases of groundedness that clearly lack philosophical content.

Firstly, consider the following way in which the natural numbers are grounded. Take some numbers, say 4, 17 and 205, and compute their sum, $4 + 17 + 205 = 226$. Thus, we have given a way of generating a natural number from some others, and a way of viewing 226 as grounded in 4, 17 and 205 (recall also figure 3 on p. 18). Of course, this case of groundedness is not interesting. This is not to say that sums are uninteresting. They may well be for pupils in primary school who have a particular leaning towards basic arithmetic. However, no point is made by calling 226 *grounded* in 4, 17 and 205.

Contrast the vacuity of sum groundedness with the case of the ordinals, that are grounded by Cantor's number generator (section 2.6). The generation of transfinite ordinals from the finite ones plays an important role in Cantor's case for the actual infinite, put forward in his 1883 *Grundlagen*. In particular, he writes that his principles of generation contribute to providing the new numbers with 'the same [...]

7 More precisely, Friedman and Sheard show the consistency of what has become known as the theory FS, see also [Halbach, 2011, §14.3].

objective reality as the earlier ones [i.e. the natural numbers]' [Cantor, 1883, p. 911]. Cantor's principles are captured by the generators **C1** and **C2** (definition 11 on p. 38). On this reading, Cantor thus ascribes metaphysical significance to the generators **C1** and **C2**. The generation of sums, in contrast, is not philosophically significant.

It may be thought that the sum generator is deficient because it is not deterministic (recall definition 6). Of course, 226 is the sum of many distinct collections of numbers. However, being deterministic is neither sufficient nor necessary for a generator to be philosophically significant. For one, the logic generators of Kripkean groundedness are not deterministic. For another, the truth generator **F** of the previous section is deterministic, but arguably not philosophically relevant.

At any rate, it is easy to cook up deterministic generators. My second example of a clearly insignificant case of groundedness is one such. Consider arbitrary, countably many $xx$. Enumerate them: $x_0, x_1, \ldots$. Now every $x \propto xx$ is grounded in $x_0$ through the generator **E** such that $y\mathbf{E}z$ iff there is an $n$ such that $y = x_n$ and $z = x_{n+1}$. Thus, $z$ is generated from $y$ if $y$ precedes $z$ in the enumeration. Since it, however, is completely arbitrary, so is **E**-groundedness of $z$ in $y$. Note that **E** is deterministic: it is exactly $x_n$ from which we generate $x_{n+1}$.

However, it is absurd to assume that this case of groundedness has philosophical significance. For one, we may begin to enumerate $xx$ at any arbitrary $y$ among them. That is, for every $y$ of $xx$ we may choose an enumeration such that $x_0 = y$. Therefore, for every $y \propto xx$ there is a generator $\daleth$ such that every $x \propto xx$ is $\daleth$-grounded in $y$. Every one of $xx$ is somehow grounded in each of them. Even if there were philosophical reasons to single out a specific $y \propto xx$ as the ground, these reasons could not lie in the general notion of groundedness but would have to be external to it.

For another, the observation may be strengthened. Any two objects whatsoever are some things $xx$, and indeed countably many. Therefore, for any two objects $x$ and $y$ whatsoever, there is a generator by which $x$ is grounded in $y$ (counting from $y$ to $x$), as well as a generator to ground $y$ in $x$ (counting from $x$ to $y$).

Thus, I have given a recipe how to ground anything, in anything. This shows that the general formal concept of groundedness from chapter 2 is excessively weak: everything is grounded somehow in anything. However, not everything is philosophically significant, fortunately so, as otherwise philosophical inquiry would be even more difficult than it already is. Hence, it cannot by itself be philosophically significant if some things satisfy the general concept.

Nonetheless, the cases of groundedness I discussed in the previous chapters, such as the iterative conception of set, or Kripke's theories of truth, have philosophical content. It is not accounted for by the general theory of chapter 2. Therefore, the theory needs to be supplemented by an account as to why certain cases of groundedness have philosophical content. In the next chapter, I will outline such an account.

## 5.5 conclusion

In this chapter I asked for the philosophical significance of groundedness and argued that this question does not have a simple answer. I considered Thomas Forster's recent case for the philosophical significance of Church-Oswald model construction, and showed that none of his arguments is conclusive. Then, I gave further example of groundedness whose philosophical significance is questionable, and ended by pointing out a simple way of viewing anything as grounded in anything. I conclude that the concept of groundedness by itself, and in general, does not ensure the grounded also to be philosophically significant. Yet, many case of groundedness have found continuing interest by philosophers, among them Kripke's concept of semantic groundedness (chapter 3), and the groundedness of sets (§ 2.7). In the remainder of this thesis, I will develop an account of why this is so, the first step towards which is taken in the next chapter.

# 6

# CONSTITUTION AND THE ITERATIVE CONCEPTION OF SETS

## 6.1 INTRODUCTION

In this chapter I will take first steps towards an account of the philosophical significance of groundedness. For this, I return to the paradigm cases presented in previous chapters. This time, I will look more closely at their philosophical content, in order to answer the question: what renders them philosophically significant? For the time being, I focus on the iterative conception of sets (section 2.7 above). Not only is it a pleasingly simple instance of groundedness, it also is arguably most extensively discussed among philosophers.[1] In particular, its philosophical content has been debated (Parsons [1977]; Potter [2004]; Incurvati [2012]). It is reasonable to hope that these discussions shed some light on how to account for the significance of the general concept, as indeed they will.

## 6.2 THE PHILOSOPHICAL CONTENT OF THE ITERATIVE CONCEPTION OF SETS

I continue working within the formal framework of chapter 2. In particular, as the iterative conception of set I understand the view that sets are obtained by iterated application of the set generator **S**, where $xx\mathbf{S}y$ iff $xx$ are the elements of $y$ (see §2.7).[2] Recall that to be grounded in nothing is not to be ungrounded, but to be generated, directly or indirectly, from nothing (p. 27). The pure sets are **S**-grounded in nothing: the first set generated is the empty set. As usual, I focus on pure sets and from now on always mean pure sets when I use 'set'.

I have already touched on the philosophical content of this view. Above, I contrasted **S**-groundedness with Forster's **ƧS**-groundedness. In particular, I observed the following difference (p. 95). Firstly, the standard set generator **S** is well glossed by saying that a set is constituted from its elements. Thus, we connect **S**-groundedness with an informal notion of *constitution*. **S** generates sets from precisely those things that constitute it. Secondly, contrary to what Forster claims, this notion of constitution is not exhausted by the extensionality of sets.

Admittedly, this informal gloss on **S**-groundedness is as imprecise as the notion of constitution involved. Nonetheless, it is usually taken as the starting point when philosophers ask for the content of the

1 Of course, Kripke's theory of truth is also widely appreciated. However, the philosophical content of semantical groundedness is seldom discussed at another than the intuitive level already found in Kripke.

2 It is subject to ongoing debate just how much of axiomatic set theory ZFC is naturally motivated from the iterative conception. Some authors have argued that the conception has difficulties accounting for Infinity and Replacement (Boolos [1998]), others disagree (Shoenfield [1977]).

iterative conception of sets.[3] The challenge is to explicate the relevant notion of constitution.

An early, influential discussion of different attempts at an explication is found in Parsons [1977]. Firstly, he examines an intuitionist understanding [§2]. According to it, a set is literally constructed from its elements. Constructed by whom? Orthodox intuitionism holds that a set is constructed by [Parsons, 1977, p. 339, my emphasis]

> [...] an idealized *finite* mind which is located at some point in time [...]

However, this approach is quickly seen to fail since it cannot account for infinite sets.

Michael Potter, in a recent discussion [2004, §3.2] which summarizes nicely much of Parsons' 1977 contribution, suggests the following response on behalf of the intuitionist. Countably infinite sets may be viewed as constructed by an idealized finite subject, if we let her carry out supertasks, that is [Potter, 2004, p. 37]

> [...] tasks which can be performed an infinite number of times in a finite period by the device of speeding up progressively [...]

A countable set thus can be viewed as constructed from them in a finite amount of time, by an idealized subject who has added the first element after one second, the second element after $1^1/_2$ seconds, the third one after $1^1/_4$ seconds and so on through all the elements of the set. After two seconds, the thought goes, she will have completed this supertask and constructed the set.

However, it has been debated whether an intuitionist may allow for constructions carried out as supertasks (Weyl [1949]). There is reason to believe that doing so would conflict with the intuitionist's rejection of the actual infinite. Moreover, as Potter remarks (ibid.), even if the concept of supertasks were available to account for the construction of countable sets, it could not help us to understand how *uncountable* sets are constructed from their elements. Therefore, a set is constituted from its elements not in the sense that it is constructed from them.

Having concluded that the constitution of a set from its elements cannot be understood as its construction from them, Parsons develops an alternative account. He proposes to understand constitution in *modal* terms. A set is constituted from its elements in the sense that it could not exist without them. Helping ourselves to possibilist quantification we may regiment this thought further: for every set, necessarily, it exists only if each of its elements exists. However, the

3 Thus, Parsons writes that '[...] one can state [...] what is essential to the 'iterative' conception: sets form a well-founded hierarchy in which the elements of a set precede the set itself' [1977, p. 336]. See also [Wang, 1977, p. 310], [Shoenfield, 1977, p. 321] and [Potter, 2004, p. 36].

modal operator, '$\Box$' as a symbol, can be used in various ways. Therefore, a modal account of how a set is constituted from its elements, is only useful if the modality at work is explicated. In his 1977 article, Parsons does not specify further his claim, that a set could not have existed without its elements. He does so, however, in his 1983.

On the one hand, Parsons provides an argument that the modality of "a set could not exist without there being its elements" is not metaphysical modality. Metaphysically, all pure sets exist necessarily. In order to account for a set being constituted from its elements, however, it is essential that a set is *contingent* on its elements,[Parsons, 1983, p. 327]

> [...] since when the elements are given the set is initially given only *in potentia*.

On the other hand, Parsons outlines a positive account how else to understand the modality, if not as metaphysical [Parsons, 1983, p. 316].

> In saying that a multiplicity of objects can constitute a set, I mean that they can do so without changing anything at "lower" levels, that is, without changing the structure of the individuals or of the sets that might have entered into the constitution of the objects making up the multiplicity in question. It is this strong possibility that the modal operator [...] is meant to express.

It is helpful to draw an analogy with a modality that we know better.[4] While it is physically necessary that I do not leap Senate House, the laws of physics do not have to change for me to jump on this table. It is in this sense that I can jump on this table while I cannot leap Senate House. Analogously, some things $xx$ do not have to change for their set $y$ to be formed. In this sense, $xx$ *can* constitute their set.

Parsons suggests one way of modelling this modality in terms of possible worlds. We may analyse the physical necessity that I do not leap Senate House by saying that I do not do so at every world where the laws of physics hold. Analogously, we may paraphrase "necessarily, there is the set $y$" as "at every stage higher up in the cumulative hierarchy there is $y$". On this basis, Parsons glosses necessity as being '[...] true "from there on" [...]' [Parsons, 1983, p. 317].

However, we cannot understand the modality of set constitution in terms of the cumulative hierarchy, if our goal is to explain why **S**-groundedness is philosophically significant. Since, the sets of the cumulative hierarchy just are all and only the **S**-grounded pure sets. Thus, explicating the modality as suggested by these remarks of Parsons would render our attempt at explanation circular. The philosophical content of **S**-groundedness is that a set could not exist without

4 Note, however, that what follows is a charitable reconstruction of Parsons' remarks.

its elements, but all we have in order to understand this 'could', is **S**-groundedness itself.

Fortunately, Parsons has more to say about how the modality of set constitution is to be understood. Later in his 1983 article [p. 328f.], he proposes to understand the constitution of a set from its elements as a modality distinct from, but related to metaphysical modality in that both specify, albeit in different ways, a general mathematical modality.

This notion of mathematical modality is not developed in detail, and may be found insufficiently clear. For the purpose of explicating the modality of set constitution, however, it suffices to note that in this general sense of possibility, mathematical entities are fully contingent. They do not all necessitate one another, as they do in the more specific case of metaphysical modality. Thus, Parsons' notion of mathematical modality allows us to speak of it being possible that some, but not all, sets exist.

The modality of set constitution is then viewed as a specification of this general mathematical modality. From a world $w$ with some sets $xx$, all and only those worlds are accessible at which each of $xx$ has just the elements that it has at $w$. In other words, $v$ accesses $w$ if and only if $w$ end-extends $v$, with respect to the relation of set elementhood $\in$. That is, if $x$ is an element of $y$ at $v$ then $x \in y$, too, at $w$. On this modality, elementhood becomes *rigid* in the precise sense that if $x \in y$ then necessarily so, and if $x \notin y$ then necessarily so, too [Parsons, 1983, p. 209].

We have thus been provided with an explication of the notion of modality in terms of which Parsons proposes to understand the constitution of a set from its elements. Moreover, this explanation does not refer directly to **S**-groundedness. Has Parsons thus succeeded and explained the philosophical content of the iterative conception of sets? I do not think so. Above, I have found that Parsons' first account of the modality of set constitution made us attempt to explain the significance of **S**-groundedness in terms of **S**-groundedness. I think that the revised explication of the previous section also leads us into a circle, as follows.

Our starting point was that **S**-groundedness is philosophically significant because the generator **S** captures the constitution of a set from its elements (cf. p. 95). This constitution Parsons now invites us to understand modally: a set *could* not exist without its element. The relevant modality, however, is explicated in terms of one world accessing another just in case the latter end-extends the former. A set $x$ is constituted from its elements $yy$, the proposal goes, in the sense that in every situation such that sets have at least the elements which they actually have, if $x$ exists so do $yy$. The significance of **S**-groundedness is that **S** captures this modal relation.

Why, however, do end-extensions matter, and not those situations in which some set does not have all elements which it actually has? Note that this distinction is drawn in terms of the elementhood relation. Relevant are those situations that respect which elements a set actually has. That is, from what a set is actually **S**-generated. The distinction between which possible situations matter and which do not, thus hinges on actual **S**-generation. Therefore, to explicate the modality to which Parsons reduces the notion of constitution, we need to resort to **S**-grounding; and to accept his case for the significance of **S**-groundedness we need to accept that actual **S**-generation matters philosophically. Similarly to before, we end up trying to account for the significance of **S**-groundedness in terms of **S**-grounding.

Maybe my reading of Parsons' 1983 article is tendentious. His remarks there may not be intended as explicating the modality which is at work in his modal account of set constitution. Here is an alternative, less demanding reading. Having earlier paraphrased the priority of elements over their set modally (a set *could not* exist without its elements existing), Parsons seeks to give his reader a better sense of how the modal vocabulary is to be understood. For this, he connects it with the well-understood concept of end-extension. Crucially, though, the modality of set constitution is not to be defined in these terms. Therefore, my circularity charge from above does not apply.

However, if this less demanding reading is appropriate and Parsons does not intend to explicate the modality, then we are left with an account of set constitution in terms of a primitive *sui generis* set modality. In this case, the question arises what advantage is gained over an iterative conception based on a primitive notion of constitution, as e.g. considered in Potter [2004].

In sum, Parsons accounts for set constitution in modal terms, but either the relevant modality is as much a primitive as set constitution itself is for other authors, or Parsons proposes an explication which relies on the concept of **S**-grounding, such that the resulting conception does not account for the significance of **S**-groundedness.[5]

## 6.3 TAKING CONSTITUTION SERIOUSLY

The formal concept of well-founded, that is **S**-grounded, sets has received much attention from philosophers. Often, it is motivated from the thought that a set is constituted from its elements or, to quote Boolos once more, (1971, p. 216)

> [...] the elements of a set are "prior" to it.

In the previous section, I have discussed Parsons' approach of explicating this intuitive idea of constitution in modal terms. Which notion

5 I will return to the connection between groundedness and modal logic in chapter 9.

of modality is right for this task? I have found that only a primitive notion of *sui generis* set-theoretic modality appears promising.

In view of this, however, we may as well return to our point of departure, and take the notion of constitution as the *primitive* of the iterative conception of set. We need not to abandon the modal approach. We may still express constitution in modal terms.[6] However, we no longer attempt to reduce constitution to some modality.

This does not mean we must give up hope to understand the priority of elements over their set. There are other ways of conveying a philosophical notion than its reduction. Examples can help us to realize our pre-theoretic grasp of it. This intuitive understanding we can then explicate by formal principles.

Consider the following English sentences.

(1) Truth and reason constitute that intellectual gold that defies destruction.

(2) Switzerland is constituted from its 26 cantons.

(3) The meaning of '+' is constituted from how this symbol is used.

These sentences are grammatical and express statements. In fact, (1) is given as an example in the 1913 Webster dictionary entry for "constitute". Speakers of English make claims like (2).[7] (3) and claims of similar form are made frequently in philosophical contexts. Philosophers accept or reject them, based on philosophical considerations. Hence, "constitute" has meaning, with a specifically philosophical aspect to it. At any rate, we understand it – at least to the extent of a pre-theoretic grasp.

In addition, we can characterize the *formal* properties of constitution. For this purpose I choose a plural meta-language. In this framework we can formalize constitution as a relation that takes at its first place singular as well as plural terms. An object thus can be constituted from a single or from several objects.

Firstly, constitution is *unique* on its left as well as on its right hand side. If $yy$ constitute $x$, and $zz$ constitute $x$, then $yy$ are $zz$. Similarly, if $yy$ constitute $x$, and $yy$ constitute $y$, then $x = y$. Uniqueness on the left has two desirable consequences. On the one hand, it renders constitution non-monotone: if $yy$ constitute $x$ then no $zz$ properly extending $yy$ can be said to constitute it. Intuitively, all constituents matter, and all that matters are the constituents.

On the other hand, uniqueness on the left renders constitution immediate: if $yy$ constitute $x$ then they do so directly, and not via some

6 In particular, the modal set theories examined very recently in Studd [2013] and Linnebo [2013] can be viewed as expressing the order of sets by constitution, see chapter 9.

7 A google search of the phrase "constituted from" provides plenty similar examples (31200 hits on October 7th, 2013).

other constituents of it. Of course, my choice of such a notion is guided by what I propose to apply it to. The generator **S** captures the direct step from some things to their set.

Uniqueness on the right is a strong assumption. However, it makes great sense for the constitution of a set from its elements. After all, sets are extensional and no two sets have the same elements, thus are constituted from the same things.

If $yy$ constitute $x$ we say that a $z$ among them *partially* constitutes $y$. Partial constitution is still immediate, but we may consider its transitive closure, *partial mediate* constitution. Note that if we understand **S** as full constitution then a statement $x \in y$ in the language of set theory expresses the partial constitution of $y$ by $x$. Thus, partial mediate constitution corresponds to one set standing in the transitive closure of another.

Secondly, constitution is *non-circular*. There is no sequence of objects $x_0, \ldots, x_n$ such that for every $i$ less than $n$, $x_{i+1}$ partially constitutes $x_i$, and $x_0 = x_n$. In particular, no $x$ partially constitutes itself.

Having described the intuitive notion of constitution by formal principles, I can render precise one way in which **S**-groundedness is connected to the philosophical notion of constitution. The relation between an **S**-grounded set and the sets from which it is **S**-generated, satisfies the formal principles of constitution. This is not at all a deep insight. To be **S**-grounded is to be a well-founded set of the cumulative hierarchy (proposition 4 on p. 41). Since partial **S**-generation is elementhood, non-circularity of course holds of the **S**-grounded sets. And, I already pointed out that by the extensionality of a set, the relation that its elements bear to it, satisfies uniqueness. Thus, the set generator **S** provides a simple model for constitution as characterized by the principles above.

On this basis, the connection between **S**-groundedness and constitution is well viewed as that between a philosophical idea and a formal model of it. Such a connection provides the model with philosophical significance, as examination of it may elucidate the philosophical idea. This is a common enough situation. Consider possible worlds semantics and its relation to the metaphysics of modality. These are distinct enterprises. Nonetheless, study of worlds semantics has had tremendous impact on metaphysics (Lewis [1986]; Kripke [1980]). It is philosophically significant. Similarly, **S**-groundedness provides a formal representation of the idea that a set is constituted from its elements, and as such it has significance.

Of course, a small set of principles as I gave for constitution, cannot pin down one concept. Even if combined with examples such as (1) to (3), they allow for different understandings of "constitute" and its inflections. At best, therefore, I have characterized a family of notions each of which is an equally good candidate for the philosophical content of **S**-groundedness.

In fact, what I have said so far is *prima facie* compatible with various philosophical notions traded under the labels of ontological priority, dependence or fundamentality. For example, consider the relation of essential dependence that $x$ bears to $y$ if $x$ is essentially such that it exists only if $y$ exists [Fine, 1995b, p. 273]. Constitution may be understood as the relation that is born to some thing by what it immediately essentially depends on. This reading appears compelling – after all, Kit Fine famously motivates his account of essential dependence from the example that a singleton depends on its element. This case of dependence is widely accepted (for a survey of the area consult Correia [2008]). Hence, viewing the philosophical significance of the set generator **S** in terms of essential dependence would connect the formal concept of well-foundedness with an established idea from metaphysics.

Fine's recent work on essential dependence was to some extent anticipated by Husserl's use of and reflection on a notion of *foundation* 2001 (see the discussion in Fine [1995a]; Correia [2004]). On a salient interpretation, Husserl distinguishes moments (roughly, tropes) from pieces (roughly, parts) as follows. A moment depends on the datum which has it, while a piece does not stand in the same relation of dependence to the whole. Now, the relevant notion of dependence is *foundation*. A moment is founded on its datum while a piece is not founded on any whole that it is a piece of. Although details are subject to controversy, it is safe to say that foundation is irreflexive: nothing is founded on itself. Now, Husserl has a notion of immediate as well as mediate foundation as its transitive closure [Husserl, 2001, §16]. Therefore, foundation must be subject to a principle of non-circularity, as I gave it above for constitution.

In Husserl's system, foundation plays the role of a general order of things which is not mereological. Foundation is that relation of ontological priority which is not the priority of parts to their whole. While the subset relation has often been likened to mereological parthood, it is widely agreed that the relation between a set and its elements cannot be understood as that between a whole and its parts [Lewis, 1991, §1.3]. Thus, if we seek a metaphysical notion to explicate the priority of elements over their set, Husserl's foundation appears a natural candidate.

To give one more candidate notion of constitution, let me return from Husserl to the contemporary debate and briefly rehearse Jonathan Schaffer's recent case for what he calls the neo-Aristotelian approach to metaphysics [2009]. It is based on a primitive relation that Schaffer calls 'grounding' [Schaffer, 2009, §3.2]. Unfortunately, Schaffer's use of this term interferes with the terminology of chapter 2, as well as with that of a certain tradition in metaphysics to which I will turn in the next chapter. However, Schaffer also speaks of *priority in nature*, and I follow him.[8]

Schaffer goes about conveying his notion of priority in nature in much the same way in which I characterized constitution. He gives examples, and formal principles. Formally, Schaffer's priority is irreflexive, transitive and asymmetric. It is a relation of *partial* priority, in that there are cases of two distinct things both being prior to the same object. Thus, priority in nature satisfies the principles of partial, mediate constitution I gave above. Schaffer's examples also make his notion a suitable candidate for constitution. In fact, it is the priority of the element to its singleton set that serves Schaffer as one of his paradigms of priority in nature [Schaffer, 2009, p. 375].

In Fine, Husserl as well as in Schaffer I have found philosophical notions that may play the role of constitution in an account of the philosophical significance of **S**-groundedness. On further investigation, they may condense into one and the same idea, or fragment further into more specific notions. At any rate, what this short excursion into the literature shows is that under the label of constitution I have not characterized a definite concept, but rather a family of cognate notions.

However, in their vicinity we also find philosophical notions that do not satisfy my characterization of constitution. One such non-example is the relation of definitional priority that Fine compares with essential dependence [Fine, 1995b, §2]. Recall the notion of real definition as opposed to nominal definition. Roughly, one thing is definitionally prior to some other if the real definition of the one involves the other. Now, however details are spelt out, it appears possible that the same plurality of things is involved in the real definition of distinct things. Thus, definitional priority is not unique on the right-hand side and therefore does not satisfy how I described constitution.

Nonetheless, it cannot be denied that the characterization of constitution given above may fit several cognate notions rather than one distinguished philosophical concept. However, for my present purpose it suffices that **S** models some notion of constitution. In chapter 5 I have found that the general formal concept of groundedness from chapter 2 does not account for the attraction certain cases of groundedness have to philosophers. It needs to be supplemented by an account of the philosophical content of these cases. In the present chapter, I focused on the groundedness of sets, and made the following proposal. The philosophical content of the generator **S** is that it tracks a notion of ontological constitution, as described by the examples and principles above. If, as it seems, there are many such notions then the philosophical content of **S**-groundedness may be understood in multiple

8 See Sider 2011, p. 192 on the difference between Schaffer's priority in nature and grounding as an *in-virtue-of* relation as in chapter 7.

ways. Each of these ways, however, would be ways in which it has philosophical significance. To show this, however, is what my goal has been. The philosophical content of the groundedness of sets may be diverse and plentiful, but such riches would be no embarrassment.

## 6.4 CONCLUSION

In this chapter, I presented an understanding of the iterative conception of set as combining the view of sets as **S**-grounded in the sense of chapter 2, with the view that this *set-of* generator **S** captures the ontological immediate, full constitution of a set from its elements. This latter notion of ontological priority is philosophical.

Hence, for one specific, albeit prominent, case of groundedness I have identified one account of its philosophically significance. Doing so, I have partially answered the challenge from chapter 5. However, more needs to be done. In particular, an account is needed of the philosophical significance of *semantic* groundedness (chapter 3).

# 7 TRUTH *IN VIRTUE OF*

## 7.1 INTRODUCTION

This chapter introduces a philosophical notion in terms of which I will then, in the next chapter, propose an account of why semantic groundedness is significant. I will explain the idea of one truth holding *in virtue of* others.

Recent years have seen an increased interest in this notion, and a certain consensus has emerged to speak of *ground* or *grounding* [Fine, 2001; Schaffer, 2009; Fine, 2012b]. Thus, authors have found it helpful to use the same label, or a label of the same semantic field, for the in-virtue-of relation as Kripke chose for his model construction. Following Kripke, I have used 'groundedness' for the class of model constructions that are the subject of the present investigation. On the one hand, I am intrigued by this terminological coincidence, and feel tempted to believe that it provides at least motivation for connecting the two areas. On the other hand, I do not want to fall for this temptation. Doing so would provoke the objection that I equivocate. Therefore, I will deviate from the current literature to the extent that I will not use the term 'ground' for the in-virtue-of relation.

The idea of one truth holding in virtue of another has a venerable tradition. It can be traced back to Aristotle's notion of *why*-proofs, in contrast with mere *that*-proofs [Aristotle, 2006, 1051b]. The first extensive discussion of the in-virtue-of relation, however, is to my knowledge found in Bernard Bolzano's *Wissenschaftslehre* of 1837 (henceforth referred to by 'WL'). I will therefore begin with a concise exposition of Bolzano's theory.

## 7.2 BOLZANO

For Bolzano, the in-virtue-of relation, 'Abfolge' in the German of his writings, is what holds truths together, and brings them into order. Accordingly, I first look at Bolzano's understanding of propositions ('Sätze').[1]

In contemporary mainstream philosophy, a proposition is what is said by an indicative sentence. For example, the sentence "Snow is white" expresses the proposition that snow is white. It is safe to understand Bolzano as working with a close kin of this notion. Importantly, for Bolzano a proposition is not located in space nor time. In this precise sense, he takes propositions to be *abstract* entities.

---

1 As to notation, I will use capital Roman letters from the beginning of the alphabet as variables for propositions. Small letters from the beginning of the alphabet, mostly 'b' and 'c', will range over Bolzanian ideas. The expression '$A(c/b)$', finally, will denote the proposition that differs from $A$ only in that everywhere where $A$ involves $b$, $A(c/b)$ involves $c$.

Further, he takes a proposition to contain *ideas*, which themselves are composed of simple ideas.[2] In fact, a proposition is well viewed as itself being *composed* of simple ideas, in the precise sense that A is B just in case that A is composed of the same simple ideas as B, in the same manner. Nonetheless, it is important to keep in mind that for Bolzano, the concept of a proposition is fundamental, and that the notion of an idea is understood in terms of it, as a component.

True propositions are ordered by the relation of one truth holding *in virtue of* others. For a long time, scholars have found this part of Bolzano's oeuvre 'obscure' [Berg, 1962, p. 151]. Bolzano motivates his theory of *Abfolge* from examples of the following kind [WL §198].[3]

(1) It is warmer in Palermo than it is in New York.

(2) The thermometer stands higher in Palermo than it does in New York.

Both propositions are true but (2) is true in virtue of the truth of (1) and not vice versa. The asymmetry between the truth of (1) and the truth of (2) cannot be captured by Bolzanian *derivability* (roughly, logical consequence extended by material implication): (1) can be derived from (2). Therefore, a stronger concept is needed: (1) is true *in virtue of* the truth of (2).

If the truth that A holds in virtue of it being true that B, then it is the case that B *because* it is the case that A. The in-virtue-of relation is a notion of objective explanation. However, it must not be conflated with epistemic notions, such as justification. For one, *in-virtue-of* concerns how propositions, that do not have spatio-temporal location, are ordered independently of any subject. For another, justification fails to respect the asymmetry between the truths (1) and (2). If you know that the thermometer stands higher in Palermo than it does in New York, then you are justified in believing that it is warmer in Palermo than it is in New York.

Bolzano discusses whether the in-virtue-of relation can be defined in terms of derivability, and possibly other notions [WL §200]; his conclusion is that this cannot be done. Therefore, the in-virtue-of relation is officially a primitive concept and Bolzano characterizes it by a system of principles. Note that Bolzano thus characterizes the in-virtue-of relation in much the same manner as constitution has been characterized in the previous chapter (§ 6.3), by examples and principles. Bolzano's in-virtue-of relation thus presents itself as a similar kind of philosophical notion as constitution. This similarity provides

2 What does Bolzano mean by 'idea' ('Vorstellung')? For present purposes, it suffices to focus on three components of Bolzano's theory. Firstly, an idea, too, is abstract, Secondly, it may have an extension of spatio-temporal objects that fall under it. Finally, as indicated above, complex ideas are individuated by their composition from simple ideas.

3 Further examples are found in [WL §§ 162.1, 201]

first support for my analogy between how set groundedness receives philosophical significance from the notion of constitution, and an understanding of semantic groundedness based on the in-virtue-of relation.

I now turn to present the principles which Bolzano puts forward for his in-virtue-of relation. *Abfolge* holds between single or collections of propositions. His relevant use of "collection" is well viewed as anticipating how logicians today have come to use the term 'plurality', as convenient but inessential shorthand for *plural* reference to some things.[4] However, Bolzano assumes a collection always to be non-empty.[5] Accordingly, I will for the rest of this section equally disregard the empty plurality.

I will use Greek capital letters ('$\Gamma$', '$\Delta$',...) as variables ranging over pluralities of propositions, and the symbol '$\lhd$' for Bolzano's in-virtue-of relation such that '$\Gamma \lhd A$' reads: it is true that $A$ in virtue of the truths $\Gamma$.

The first principle set forth by Bolzano is that only true propositions stand in the in-virtue-of relation [WL §203].

FACTIVITY If $A_0, A_1, \ldots \lhd B_0, B_1, \ldots$, then $A_0, A_1, ..., B_0, B_1 \ldots$

We therefore know that

(3) Francesca is male.

does not hold in virtue of

(4) Every sister is male.

If the truth that $A$ holds in virtue of the truth that $B$, then the latter is *why* it is the case that $A$; the fact that $B$ explains the fact that $A$. The sense of explanation at work here is objective and exhaustive. This allows us to draw two conclusions about the formal properties of *Abfolge*. Firstly, what a proposition holds in virtue of does not involve this proposition itself, neither directly or indirectly [WL §§204, 218].[6]

NON-CIRCULARITY There is no chain $A_0, ..., A_n$ such that for every $i < n$, $A_i$ is among some $\Gamma$ such that $\Gamma \lhd A_{i+1}$, and $A_0 = A_n$.

Finally, the truths which $A$ holds in virtue of are *unique* [WL §206].

UNIQUENESS If $\Gamma \lhd \Delta$ and $E \lhd \Delta$ then $\Gamma$=E.

On the one hand, this implies that if $A$ holds in virtue of $\Gamma$, then it is not grounded in $\Gamma$ together with arbitrary other truths. Thus, it is ensured that every truth among $\Gamma$ matters for the truth $A$. In other

4 However, Bolzano uses 'collection' also in other ways [Simons, 1997]
5 In one passage Bolzano may also be read as suggesting that the relata of the in-virtue-of relation are always *finite* pluralities of propositions [WL §199].
6 This formulation of non-circularity uses the concept of a in-virtue-of *chain* due to Rumberg [2013].

words, *Abfolge* is non-monotone. On the other hand, the principle of uniqueness means that the truths that a proposition is grounded in, are its *complete* grounds. This captures our pre-theoretic idea of grounding as a relation of exhaustive explanation.

Note, however, that Bolzano's formal principles do not include uniqueness on the right-hand side. That is, it is not assumed that if $\Gamma \lhd A$ and $\Gamma \lhd B$ then $A = B$.

These principles describe the relation of *Abfolge* formally. For example, from (Uniqueness) we know that *if* the truth

(5) Michael is a son of Vito.

holds in virtue of the truth that

(6) Vito is Michael's male parent.

then it is not the case that (5) holds in virtue of the truth that

(7) Sonny, Fredo and Michael are Vito's sons.

We would like to know more. Does (5) in fact hold in virtue of (6)? More generally, what cases of *Abfolge* are there? Bolzano gives examples, but not many general principles. However, there is one prominent exception, which will become central in the next chapter [WL §205.1, my translation].

> Let $A$ be any truth: then the truth "that the proposition $A$ is true," is a proper consequence of it; and this consequence does certainly not need grounding in any other truth than $A$ alone, which therefore constitutes its full ground.

Thus, for every truth $A$, it is true that $A$ in virtue of it being the case that $A$, $A \lhd T(A)$. By the same principle we have that it is true that $A$ is true in virtue of the proposition that $A$ is true itself. In symbols: $T(A) \lhd T(T(A))$.

For Bolzano, truth is an idea. Since propositions are identified by how they are built up from which ideas, the proposition that $A$ therefore is not identical to the proposition that $A$ is true. Hence, *uniqueness* ensures it not to be the case that $A \lhd T(T(A))$.

Generally, *Abfolge* is a notion of *complete*, *immediate* objective explanation. From it, *partial* such explanation is defined easily [WL §198]: $A$ holds *partially* in virtue of the truth $B$ if there are some $\Delta$ such that $B$ is one of $\Delta$, and $A \lhd \Delta$.

Bolzano does not stop at the relation between a truth and what it holds in virtue of immediately. He analyzes the order that it imposes on true propositions [WL §216].

> If someone starting from a given truth $M$ asks for its ground, and if finding this in [...] the truths $[A, B, C \ldots]$

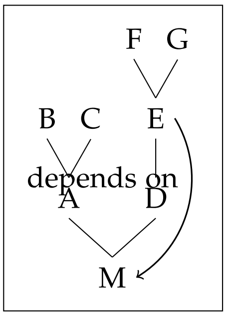


Figure 14: Ascension from M to its auxiliary truths

> he continues to ask for the [. . . ] grounds, which [. . . ] these have, and keeps doing so as long as grounds can be given: then I call this *ascension from consequence to grounds*.

If, ascending from M to what it holds in virtue of, we arrive at some truth A, then M is said to *depend* on A.

The next section brings us back to contemporary philosophy. It is a focused survey on recent work by Kit Fine, in which he re-examines Bolzano's notion of one truth holding in virtue of others, provides a rigorous formal system to reason about it and puts it to new use in contemporary metaphysics, thus showing its philosophical significance.

## 7.3 fine

In his 2012 article *The Pure Logic of Ground*, Kit Fine revives many of Bolzano's ideas and develops them in a modern framework. In particular, he presents a formal system to reason about the in-virtue-of relation, or rather four such relations, each capturing one specific aspect of the pre-theoretic notion.

Firstly, Fine distinguishes between a weak and a strict sense of 'in virtue of'. On the one hand, to say that C holds weakly in virtue of A, B, is to say that for it to be the case that C is for it to be the case that A, B, . . . [Fine, 2012a, p. 3]: In particular, any truth holds weakly in virtue of itself. In allowing for this reflexive in-virtue-of relation, Fine goes beyond Bolzano who did not leave space for it in his theory (see p. 122 above). Strict in-virtue-of, on the other hand, is irreflexive. Adopting a useful metaphor of Fine's, the strict in-virtue-of relation moves us '. . . down in the explanatory hierarchy', while weak in-virtue-of has moves merely 'sideways' [Fine, 2012a].

Secondly, like Bolzano Fine distinguishes between *full* and *partial* in-virtue-of. While Bolzano, however, derives partial mediate in-virtue-of from his basic notion of immediate full in-virtue-of (see above), Fine introduces both as primitives and draws the distinction intuitively, as follows. A holds fully in virtue of $\Gamma$ if $\Gamma$ are sufficient for A.

| | Strict | Weak |
|---|---|---|
| Full | $<$ | $\leqslant$ |
| Partial | $\prec$ | $\preccurlyeq$ |

Table 1: Fine's concepts of ground.

*Partial* grounds $\Gamma$ , on the other hand, merely help ground $A$: there are other $\Delta$ such that $\Gamma$ and $\Delta$ together suffice for $A$.

Fine presents his pure logic of ground as a system to derive sequents of the form "$A$ in virtue of $\Delta$". Since I intend to use Fine's system to enrich a given theory, say of truth, by the resources to speak of the in-virtue-of relation, I will transfer it to in a language $\mathcal{L}_g$ which extends some first-order language $\mathcal{L}$ by the four sentential operators of table 1. Since in earlier chapters, I use lower-case letters from the middle of the Greek alphabet as schematic variables for $\mathcal{L}$-sentences. Further, as in the previous section, I use upper-case letters from the Roman alphabet for propositions (abusing the use-mention distinction occasionally). Thus, within $\mathcal{L}_g$ the in-virtue-of relation is expressed in a way analogous to how in English it may be expressed by the connective "because". This approach allows us to make do without additional resources to speak of the fact that $A$, or the proposition that $A$ in a first-order setting.

Since we would like to express that some fact holds in virtue of multiple facts *together*, two of the four sentential operators are of *variable arity*. That is, $\phi_0 < \psi$ and $\phi_0, \phi_1 < \psi$ as well as $\phi_0, \ldots, \phi_n < \psi$, for any $n$, are sentences of the language.

This may not even be enough. We may want to to say that it is true that $\phi$ in virtue of infinitely many truths. For this, we would need to formulate the pure logic of ground in a language with formulae of infinite length. In general, it is desirable not even to impose any ordinal bound on the arity of '$<$', in which case we would need to allow for formulae of absolutely infinite length.

However, such non-standard infinitary languages are not well understood. Therefore, I will restrict my attention to a finitary language of ground. As a consequence, in the next chapter I will only be able to capture the sentential fragment of Kripke's logic generators. A quantified truth of the form $\forall x \phi$, namely, would have to be grounded in infinitely many truths $\phi(a), \phi(b)$ and the fact that $a, b, \ldots$ are all and only the things which exist.

I believe that this limitation does not undermine my case for the philosophical significance of semantic groundedness. For one, my present goal is merely to outline one such account. Future research will clarify the behaviour of infinitary systems of ground. Then, their modelling by semantic groundedness can be extrapolated from my work on the finitary fragment. For another, the dimension in which my present case will be limited is orthogonal to the connection between semantic groundedness and the in-virtue-of relation. And it is the latter that renders plausible my analogy to **S**-groundedness and answers the challenge from chapter 5. At any rate, to the best of my knowledge the details of Fine's system in an infinitary language have not yet been spelt out, and I will not attempt to answer this substantial open question here.

Now let the *pure logic of ground* be an $\mathcal{L}_g$-theory, defined as follows.

**Definition 25** (Pure Logic of Ground)**.** Let $\mathcal{L}_g$ extend some first-order language $\mathcal{L}$ by the connectives of table 1. The *pure logic of ground* ('PLG') in $\mathcal{L}_g$ then is the least set of $\mathcal{L}_g$-sentences containing the axioms of Identity and Non-Circularity (see below), and closed under the following rules.

Firstly, we have *subsumption* rules. They allow for the inference of weak in-virtue-of claims from strict such, and of partial in-virtue-of statements from full in-virtue-of.

$$\mathrm{S}\left(\frac{<}{\leqslant}\right)\ \frac{\phi_0, \phi_1, \ldots < \psi}{\phi_0, \phi_1, \ldots \leqslant \psi} \qquad \frac{\phi \prec \psi}{\phi \preccurlyeq \phi}\ \mathrm{S}\left(\frac{\prec}{\preccurlyeq}\right)$$

$$\mathrm{S}\left(\frac{<}{\prec}\right)\ \frac{\phi, \zeta_0, \zeta_1, \ldots < \psi}{\phi \prec \psi} \qquad \frac{\Gamma, \phi \leqslant \psi}{\phi \preccurlyeq \psi}\ \mathrm{S}\left(\frac{\leqslant}{\preccurlyeq}\right)$$

Secondly, three rules ensure the transitivity of partial ground.

$$\mathrm{T}\left(\frac{\preccurlyeq}{\preccurlyeq}\right)\ \frac{\phi \preccurlyeq \psi \qquad \psi \preccurlyeq \chi}{\phi \preccurlyeq \chi} \qquad \frac{\phi \preccurlyeq \psi \qquad \psi \prec \chi}{\phi \prec \chi}\ \mathrm{T}\left(\frac{\prec}{\preccurlyeq}\right) \qquad \frac{\phi \preccurlyeq \psi \qquad \psi \prec \chi}{\phi \prec \chi}\ \mathrm{T}\left(\frac{\preccurlyeq}{\prec}\right)$$

Thirdly, '$\leqslant$' obeys the following rule of Cut, the analogue of transitivity for a many-one operator.

$$\mathrm{Cut}\left(\frac{\leqslant}{\preccurlyeq}\right)\ \frac{\zeta_0, \ldots, \zeta_n \leqslant \psi_0 \qquad \xi_0, \ldots, \xi_n \leqslant \psi_1 \qquad \ldots \qquad \psi_0, \psi_1, \ldots \leqslant \phi}{\zeta_0, \ldots, \zeta_n, \xi_0, \ldots, \xi_n, \ldots \leqslant \phi}$$

Fourthly, a rule is added that allows us to infer strict full in-virtue-of from a number of partial full in-virtue-of statements, and with the statement that those partial grounds together are *weak* full grounds. So, while the subsumption rules from above allowed us to move from full in-virtue-of to weak and partial in-virtue-of, the following rule allows us to regain the strict relation. Hence the label 'reverse subsumption', or 'RS' for short.

$$\mathrm{RS}\ \frac{\phi_0, \phi_1, \ldots \leqslant \psi \qquad \phi_1 \prec \psi \qquad \phi_2 \prec \psi \qquad \ldots}{\phi_0, \phi_1, \ldots < \psi}$$

Finally, we are given axioms. Here, I deviate slightly from Fine's own presentation of his system. Recall that he formulates it as a sequent calculus, while in the language $\mathcal{L}_g$, '$\prec$' and the other symbols are sentential operators. Thus, while Fine's formalism does not allow for negated ground statements, mine does. Accordingly, while his non-circularity axiom is of the form $\frac{\phi \prec \phi}{\bot}$, I can use negation and will do so.

$$\text{Identity}\ \frac{}{\phi \leqslant \phi}\ \phi \in Sent_{\mathcal{L}} \qquad \phi \in Sent_{\mathcal{L}}\ \frac{}{\neg(\phi \prec \phi)}\ \text{Non-Circularity}$$

This completes the definition of PLG.

As in previous chapters I considered theories, e.g. the theory of classes HSK (p. 67), that are not closed under classical logic, it may be worth pointing out that the theory PLG is thoroughly classical.

It has simple models. Let $\phi_0, \dots, \phi_n$ be an enumeration of $\mathcal{L}$-sentences, and let $\phi_i < \phi_j$ iff i less than j. Let $\phi_i \leqslant \phi_j$ iff i less than or equal to j, $\phi_i \prec \phi_j$ iff $\phi_i < \phi_j$ and $\phi_i \preccurlyeq \phi_j$ iff $\phi_i \leqslant \phi_j$. Then, the subsumption rules, reverse subsumption and identity are satisfied by definition. Cut holds by the transitivity of *less than or equal to*, as does the transitivity rule for $\preccurlyeq$. The other transitivity rules hold by the transitivity of $<$ and the definition of $\leqslant$. Finally, Non-Circularity holds because $<$ is well-founded.

This observation is easily generalized. PLG holds in a model if $<$ denotes an order of $\mathcal{L}$ sentences isomorphic to the less-than relation on an initial segment of the natural numbers, in terms of which the other relations are defined as above.

Having presented his recent technical contribution, I now turn to Fine's earlier and very influential case for the philosophical significance of the in-virtue-of relation [Fine, 2001]. He proposes to discuss positions of realism in terms of the in-virtue-of relation. One prominent anti-realist position is the view that mathematical statements reduce to logic ("logicism"). Fine argues that by focusing on in-virtue-of claims, a better understanding is gained of the dispute between logicist and Platonist.

Fine points out that the anti-realist must account for the felicity of ordinary existence claims, since otherwise her position collapses into skepticism. Hence, the anti-realist needs to distinguish between two conceptions of reality. According to the *ordinary* conception, there are, say, prime numbers between 2 and 6. However, the logicist holds, this is not really the case. On the proper *metaphysical* conception, namely, there are no numbers. This metaphysical reality can be understood in two ways.

If realism about a proposition A is understood in the *factual* sense, the realist holds that A is true or false in virtue of how the world is like. Conversely, anti-realism is the view that there is no fact of the matter whether A.[7] Examples are expressivism in meta-ethics, formalism about mathematics and instrumentalism about science. In short,

7 For readability, in this section I often suppress the use-mention distinction. Thus, I use the capital Roman letters both as variables ranging over propositions, and as schematic letters standing for indicative sentences.

if reality is understood the fundamental way, anti-realism about A says that A fails to 'perspicuously represent the facts' [p. 3].

The alternative understanding of metaphysical reality is to think of it as the basis to which can be reduced what is said to be real in a merely ordinary manner [p. 8]. Thus, the realist about some proposition holds it to be irreducible, while the anti-realist holds that it can be reduced to different propositions. Logicism is anti-realism in this sense. Another example is naturalism about ethics, according to which ethical truths reduce to facts about the physical domain.

Fine now argues in considerable detail that neither factual nor reductivist anti-realism is a stable position. For the present purpose, I do not have to follow his discussion too closely. Suffice it to report Fine's final argument why any attempt to formulate non-skeptical anti-realism in terms of factuality or reduction is bound to fail [p. 11]. Anti-realism intends to be compatible with ordinary discourse. For example, naturalism about ethics must be compatible with the fact that outside of philosophical contexts, it is taken for granted that there are moral values. However, that this particular belief happens to be the ordinary view must not bear on how the anti-realist position is formulated. Consequently, anti-realism must also not be at odds with any other view that has happened not to be the ordinary opinion. Hence, non-skeptical anti-realism must be compatible with arbitrary ordinary-discourse positions. For a fair assessment of the dispute between realist and anti-realist, their positions must not be formulated in any way that incurs conflict with ordinary discourse nor with what may have happened to be the received opinion.

It is to this methodological problem that Fine proposes the in-virtue-of relation as a solution. Fine argues that the anti-realist about some truth disagrees with the realist about in virtue of what this propositions holds. Since the in-virtue-of relation is a specifically metaphysical notion, it thus provides way of adjudicating between realist and anti-realist positions, that is independent of ordinary discoure.

In a nutshell, Fine's argument goes as following. Consider the dispute between a moral realist and an expressivist. Both agree that, for example, rape is wrong. Whence the disagreement? Assume further that Jones says that rape is wrong. Again, both realist and expressivist agree that she does. However, while the realist holds that Jones says so in virtue of the fact that "wrong" refers to wrongness, the expressivist rejects this in-virtue-of claim. For him, Jones says that rape is wrong in virtue of some other truths, for example that "wrong" functions like "boo!". Thus, while moral realist and expressivist can agree that rape is wrong, thus both staying in line with ordinary discourse, the in-virtue-of relation enables them to pin down the subject matter of their dispute.

Fine considers how such disagreement about in-virtue-of cases is settled [§ 7]. For one, Fine submits that our intuitions about what holds in virtue of what are by and large reliable, and thus provide already some guidance. For another, evidence for or against a statement of the form ‘B because A’ can be found in the candidate ground A itself. The reason is that grounds are explanations, in fact explanations of superior character. A is a good candidate ground to the extent that it is the best explanation of the fact that B.

Accordingly, in-virtue-of claims should be assessed by the same standards as explanations more generally: ‘... simplicity, breadth, coherence, or non-circularity’. Fine points out that arguably, these standards vary across contexts. Explanatoriness, therefore, must be assessed in context. Accordingly, questions of in-virtue-of cannot be properly answered in isolation but only in context.

A critical discussion of Fine’s case for the in-virtue-of relation as a general tool in realism vs. anti-realism debates goes beyond the scope of the present study. Its primary subject is the concept of groundedness from philosophical logic. For my purposes, it suffices to have exposited his application of the notion, and to have it thus shown philosophically significant. Just as in the case of the iterative conception of sets the significance of philosophical notions of constitution allowed me to respond to the challenge from chapter 5, I now turn to develop an analogous defence of semantic groundedness as in chapter 3.

## 7.4 CONCLUSION

In this chapter I introduced the philosophical notion of one truth holding in virtue of others. In doing so I have achieved two ends. On the one hand, I explained the notion and provided tools to talk about it. On the other hand, I showed that it has philosophical significance.

The in-virtue-of relation has a venerable history, and presented it in chronological order. My concise summary of Bolzano’s theory gave its key properties. For example, I noted that the in-virtue-of relation is stricter than logical consequence (p. 121). I also presented methodological considerations that guide research to the present day, in particular the role of examples. Then, I turned to recent work by Kit Fine. I focused on two influential papers, and firstly presented his 2012a *Pure Logic of Ground*. Then, I turned to Fine’s 2001 case for the in-virtue-of relation as a powerful tool to carry out realism-antirealism debates in multiple areas. Fine’s investigations show that the in-virtue-of relation is philosophically significant. Thus, the present chapter has provided me with the tool to defend semantic groundedness against the challenge from chapter 5. In the next chapter, I will establish a robust connection between the in-virtue-of relation and Kripke’s concept of semantic groundedness, in particular making use of Fine’s 2012 regimentation.

# 8

# AN ITERATIVE CONCEPTION OF TRUTH

## 8.1 INTRODUCTION

Having in the previous chapter introduced the in-virtue-of relation, I will now make use of it to account for the philosophical significance of semantic groundedness (chapter 3). My proposal will be analogous to how in chapter 6 I accounted for the significance of set groundedness in terms of a philosophical notion of constitution (§ 6.3). There, I developed a view according to which the set generator **S** expresses the thought that a set is constituted from its elements. The relevant notion of constitution is not reduced along constructivist lines, or in modal terms, but taken as a primitive and characterized by means of examples and formal principles. In much the same way, the previous chapter characterized the in-virtue-of relation. In the present chapter, I will argue that semantic groundedness exemplifies this philosophical notion.

My proposal is simple. Recall the *high-resolution* characterization of semantic groundedness (pp. 50ff.). I analyzed Kripke's jump in terms of two generators: the truth generator **T**, and a logic generator, such as **W**. They are given by certain rules. Now, I argue that these rules express certain principles about what holds in virtue of what. A sentence is semantically grounded with respect to Strong Kleene logic, if and only if it is **T**-**W**-grounded in true arithmetic. Let a sentence $\phi$ be **T**-**W**-generated from $\psi, \zeta, \xi$. My proposal, at its core, is that this formal relation is philosophically significant to the extent that it is the case that $\phi$ in virtue of it being the case that $\psi$, $\zeta$ and $\xi$. In a slogan, the in-virtue-of relation is for semantic groundedness what constitution is for set groundedness.

This *iterative conception of truth* is a simple but fruitful philosophical account of semantic groundedness. Moreover, I will support it by technical results. I will show that the way *in-virtue-of* orders the truths in Kripke's least fixed point mirrors the structure of **T**-**W** priority relations. Two very recent pieces of literature give reason to expect such a connection. In his [2010], Kit Fine uses Kripke's fixed point constructions to show that certain principles of the in-virtue-of relation can be consistently combined with principles about what facts, propositions or truths there are. Fabrice Correia shows that the semantically grounded sentences stand in a relation to the true literals of the base language, which behaves formally much like the reflexive closure of the in-virtue-of relation [2013, p. 5, theorem 7.11]. My work in this chapter is of course inspired by Fine's and Correia's work.[1] At the end of this chapter, I will explain the respects in which I go beyond what they have done.

I want to propose an account of the philosophical content of semantic groundedness analogous to how I accounted for the significance of **S**-groundedness in chapter 6. However, semantic groundedness is

1 Although I learnt of Correia's insights not as early as would have been desirable.

more complicated than **S**-groundedness. The latter we saw to coincide with the concept of being a well-founded set (§2.7). Semantic groundedness is less mundane. It involves two generators, and both the truth generator **T** as well as the logic generator will require interpretation in terms of the in-virtue-of relation. It is reasonable to discuss **T** and **W** separately; I begin with the truth generator, as matters here are comparably simple.

## 8.2 truth

The truth generator **T** is given by two rules.

$$\frac{\phi}{T\ulcorner\phi\urcorner} \qquad \frac{\neg\phi}{\neg T\ulcorner\phi\urcorner}$$

In a nutshell, I will make the following case. Groundedness by the truth generator **T** is philosophically significant because it is true that A in virtue of it being the case that A.[2] Thus, the truth generator **T** gains philosophical significance from its connection to the philosophical notion discussed in the previous chapter.

I will develop my case in two steps. Firstly, I will attempt to render plausible that it is true that $\phi$ in virtue of it being the case that $\phi$, and that it is not true that $\psi$ because $\neg\psi$. To simplify matters, I will focus on the first claim, and refer to it as the *true because* claim.[3] I will show that it does philosophical work. In particular, it has been used by philosophers in order to understand better the in-virtue-of relation. On this basis, I propose to read the generator **T** as expressing this *true because* claim, thus providing the former with philosophical content. Having made this rather intuitive point, I will in a second step support it by a technical result. I will show that Fine's PLG of the previous chapter enriched by *true because*. is sound and complete with respect to the **T**-priority relations.

My first step, however, is to argue for the *true because* claim. To begin with, it seems just right to say, for example, that it is true that snow is white because snow is white. Recall from the previous chapter how Fine suggests to settle a question whether some fact holds in virtue of another (p. 128). Thus, the *true because* claim receives plausibility from arguments that to say that snow is white is to explain why it is true that snow is white, and to do so better than by any other explanation.

2 As in the previous chapter, I use upper-case letters from the Roman alphabet for propositions, abusing the use-mention distinction occasionally. Lower-case letters from the Greek alphabet, such as '$\phi$' and '$\psi$' are used as variables for object language sentences, as in the statement of the **T** rules above.

3 The present chapter is the first step towards an account of semantic groundedness in terms of the in-virtue-of relation. Accordingly, my goal is not completeness but a convincing outline. Among those details which will be worked out elsewhere is an examination of the thought that it is not true that $\phi$ because $\neg\phi$ analogous to how in the main text I examine the *true because* claim.

Now, if we are asked, why is it true that snow is white? then to say, because snow is white, is to answer the question. This assumes that to say that it is true that snow is white is not to say merely that snow is white. Accordingly, the question under consideration may well not be a common one; instead, 'why is it true that snow is white?' may usually be asked with the meaning of 'why is snow white'? Nonetheless, in philosophical contexts it is also asked with its strict meaning, not asking why snow is white, but why it is true that snow is white. And this question is answered by saying that snow is white.

However, this fact by itself does not provide evidence for the *true because* claim. We also need that no better answer can be given. How do we compare explanations? Again, it is worthwhile recalling Kit Fine's [2001] considerations (p. 128 above). One criterion is simplicity, and our candidate answer does certainly well in this respect. Since, any explanation of the truth that snow is white must involve that snow is white. In this respect, to say that snow is white is to give the simplest explanation.

Surely, more needs to be said to establish the *true because* claim. The explanation has to be assessed against other criteria than simplicity. However, what properties make an explanation stand out, depends on the domain of discourse of the proposition whose truth is explained (see, again, Fine [2001, pp. 22ff]). Fortunately, though, my goal is merely to render plausible that it is true that $A$ in virtue of it being the case that $A$. To render plausible a claim is less demanding than to establish it; I do not aim to establish the *true because* claim. My case for it is not intended to be conclusive. Therefore, I do not need to, and will not, give more argument that the best explanation why it is true that $A$ is that $A$.

Instead, I point out that the *true because* claim has an excellent pedigree. Aristotle famously writes, in book Θετα of his *Metaphysics* ($1051^b$,8f, translated by Makin 2006)[4]

> [. . . ] it is not because of our truly thinking you to be pale that you are pale, but it is rather because you are that we who say this speak the truth.

It may be contested that Aristotle here uses "because" in a sense sufficiently close to the in-virtue-of relation. Bolzano, however, is explicit that $A$ is true in virtue of the truth that $A$. I gave the key passage already on p. 123 above, but let me indulge in it once more (WL §205.1, my translation).

> Let $A$ be any truth: then the truth "that the proposition $A$ is true," is a proper consequence of it; and this consequence does certainly not need grounding in any other

4 See also Categories, Ch. 12, as cited in [Batchelor, 2010, p. 66].

truth than $A$ alone, which therefore constitutes its full ground.

The *true because* claim figures also in contemporary philosophy. In particular, it has been endorsed in the recent literature on the in-virtue-of relation itself. For example, Benjamin Schnieder [2011] deploys the following principle, which he labels simply 'Truth': for every truth '$A$', it is true that $A$ because $A$. Schnieder is explicit that here, 'because' expresses the in-virtue-of relation of my interest, such that Schnieder's principle Truth is what above I called the *true because* claim.

In fact, Schnieder not merely endorses, but uses the *true because* claim to find out things about the in-virtue-of relation. In particular, it plays a key role in his argument that the in-virtue-of relation is more finely grained than intensional operators. Assuming the principle Truth, Schnieder points out that necessarily, $A$ if and only if it is true that $A$. However, one cannot be substituted for the other *salva veritate*: since the relevant meaning of 'because' is asymmetric, the principle Truth implies that it is not because $A$ that it is true that $A$ (also recall the first conjunct of Aristotle's verdict above).

Note that I have just cited a philosopher who not only endorses the *true because* claim, which I propose as the philosophical content of the generator **T**, but also makes use of it in his investigation into the in-virtue-of relation. Elsewhere [2012, §4.4], Correia and, again, Schnieder use the *true because* claim to present a problem for truth-maker theorists who claim that it is true that snow is white because the fact that snow is white exists. Prima facie, this verdict conflicts with with the *true because* claim, or its instance: it is true that snow is white because snow is white. The natural way of reconciling this verdict ('TM') with *true because*, Correia and Schnieder submit, is by inferring TM from the in-virtue-of statement together with the verdict TM*, equally suggested by truth-maker theories: snow is white because the fact that snow is white exists.[5] Thus, if she accepts the *true because* claim, the truth-maker theorist is under pressure to likewise accept TM*. However, TM* conflicts directly with the compelling thought that the fact that snow is white exists because snow is white. Unlike TM and TM*, this thought is plausible independently of truth maker theory. Consequently, the *true because* claim enables a challenge to the truth maker theorist.

That it plays this important role for philosophical work is not a recent phenomenon. Already Bolzano made such use of it. Let me give two examples. Firstly, in the section from which I quoted above he intends to show that some truths have exactly one ground (WL §205.1). Witness every truth of the form 'It is true that $\phi$', that each holds in virtue of precisely the truth that $\phi$.

5 TM follows from the *true because* claim and TM* by the transitivity of in-virtue-of.

Secondly, in WL §214 Bolzano argues that for every $A$ another one holds in virtue of it; again, witness the truth that it is true that $A$. This thought, from which I propose to understand Kripke's truth generator **T**, therefore does work in Bolzano's investigation.

Together with my previous findings from contemporary literature, this aspect of Bolzano's work suggests one way of defending the philosophical significance of **T**-groundedness. Read the generator **T** as expressing the philosophical idea that it is true that $A$ in virtue of it being the case that $A$. Being endowed with this reading in terms of a philosophical notion, **T**-groundedness differs from vacuous cases of grounding, such as that of the sum-generator (§ 5.4). Thus, my proposal answers the challenge of chapter 5.

**T**-groundedness is significant because **T** allows us to generate sentences according to the *true because* claim. This is a simple thought, but not simplistic. In the remainder of this section, I will present a technical result, that a formal theory based on Fine's *Pure Logic of Ground* and the *true because* claim, is sound and complete with respect to the structure of **T** priority relations.

Recall that given a generator $\gimel$, we say that $x$ is $\gimel$ grounded in some $gg$ ('$gg \leqslant_{\gimel} x$') if $x$ has a $\gimel$-priority tree whose leaves are $gg$ (definition 1 and proposition 1). If we require $x$ to be generated by at least one step, we speak of it as strictly $\gimel$-grounded in $gg$ ('$gg <_{\gimel} x$'). Further, we say that $y$ $\gimel$, $gg$-depends on $x$ ('$x \prec_{\gimel,gg} y$') if $y$ has a $\gimel$-priority tree, whose height coincides with the rank of $y$, and one of whose leaves is $x$ (definition 5). Finally, I write $x \preccurlyeq_{\gimel,gg} y$ if $x \prec_{\gimel,gg} y$ or $x = y$.

The **T** priority relations satisfy the formal principles of in-virtue-of relation, in particular Fine's *Pure Logic of Grounding* ('PLG'). The reason is that a **T**-priority tree is a rather simple structure. It is a sequence of $\mathcal{L}_{\text{ta}}$-sentences $\psi_0, \ldots, \psi_n$ such that $\psi_{i+1} = T\ulcorner\psi_i\urcorner$ or $\psi_{i+1} = \neg T\ulcorner\phi\urcorner$ and $\psi = \neg\psi_i$. Thus, $<_{\mathbf{T}}$ is a strict well-ordering isomorphic to the less-than relation on a finite initial segment of the natural numbers.

This simple observation already suggests a close connection between **T** generation and the in-virtue-of relation. However, to establish such a connection to Kripke's concept of semantic groundedness, I need to be more specific. **T**-dependence $\prec_{\mathbf{T},\Sigma}$ is relative to some truths $\Sigma$ from which other truths are **T**-generated. My interest is in specific such grounds. Recall that Kripke's semantic groundedness, based on Strong Kleene logic, is **T**-**W** groundedness in those literals $\Lambda$ of the language of arithmetic which are true in the standard numbers (lemma 5 on p. 53). Consequently, it is **T**-dependence relative to $\Lambda$ that I will focus on.

However, because a **T**-priority tree is just a sequence $\psi, T\ulcorner\psi\urcorner, T\underset{\cdot}{T}\ulcorner\psi\urcorner, \ldots$, T-dependence $\prec_{\mathbf{T},\Lambda}$ in fact coincides with the relation $<_{\mathbf{T}}$ on the **T**-**W**-grounded sentences of $\mathcal{L}_{\text{ta}}$. Therefore it, too, is a strict well-ordering isomorphic to the less-than relation on a finite initial segment of the natural numbers. Hence, each $\mathbf{T}, A$-priority tree is a PLG model of precisely the simple kind described in section 7.3 (p. 127). Consequently, the relations $\leqslant_{\mathbf{T}}, <_{\mathbf{T}}, \prec_{\mathbf{T},\Lambda}$ and $\preccurlyeq_{\mathbf{T},\Lambda}$ satisfy the principles of PLG. More precisely,

**Lemma 12.** *Let $\mathrm{I}_{\boldsymbol{T}\text{-}\boldsymbol{W}}$ be the set of $\mathcal{L}_{\mathrm{ta}}$-sentences $\boldsymbol{T}$-$\boldsymbol{W}$-grounded in $\Lambda$, and let $\mathfrak{T} = (\mathrm{I}_{\boldsymbol{T}\text{-}\boldsymbol{W}}, <_{\boldsymbol{T}}, \leqslant_{\boldsymbol{T}}, \prec_{\boldsymbol{T},\Lambda}, \preccurlyeq_{\boldsymbol{T},\Lambda})$ be the structure of grounded sentences with the $\boldsymbol{T}$ priority relations. Let $\mathcal{L}_{\mathrm{gta}}$ be the language of truth extended by the grounding operators, as in definition 25, and let PLG be the least set of sentences in this language closed under Fine's rules of the pure logic of ground there defined. Then*

$$\mathfrak{T} \models \mathrm{PLG}$$

For example, $\mathfrak{T}$ satisfies the transitivity rule $\mathsf{T}(\substack{\preccurlyeq\\ \prec})$ because the $\mathcal{L}_{\mathrm{gta}}$-sentence $\phi \preccurlyeq \psi$ holds in it only if either $\phi$ and $\psi$ are the same $\mathcal{L}_{\mathrm{ta}}$ sentence in $\mathrm{I}_{\mathbf{T}\text{-}\mathbf{W}}$, or $\phi$ occurs as a vertex in one of $\psi$'s $\mathbf{T}$-priority tree whose height is $\psi$'s $\mathbf{T}$-$\Lambda$ rank. Either way, the $\mathbf{T}$ tree witnessing $\phi \prec_{\mathbf{T},\Lambda} \psi$ is, respectively can be extended to a witness of $\phi \prec_{\mathbf{T},\Lambda} \psi$, hence $\mathfrak{T} \models \phi \prec \psi$.

The formal concept of $\mathbf{T}$-generation provides a simple but precise model of grounding as characterized by Fine's pure logic. Thus, it stands to this philosophical notion much like the set generator $\mathbf{S}$ stands to the notion of constitution from section 6.3. Such a connection already provides the formal concept with some philosophical significance (p. 114).

We can do better. What I propose to take as the philosophical content of $\mathbf{T}$-groundedness is neatly expressed in the language $\mathcal{L}_{\mathrm{gta}}$, by sentences of the form.

$$\phi < \mathsf{T}\ulcorner\phi\urcorner \tag{45}$$

$$\neg\phi < \neg\mathsf{T}\ulcorner\phi\urcorner \tag{46}$$

Let us add as axioms to the $\mathcal{L}_{\mathrm{gta}}$-theory PLG every statement of the form (45), for $\phi \in \mathrm{I}_{\mathbf{T}\text{-}\mathbf{W}}$, and each instance of (46) for $\neg\phi \in \mathrm{I}_{\mathbf{T}\text{-}\mathbf{W}}$. I will refer to these sentences as the *truth* axioms. The resulting theory GT, of the *grounds* of *truth*, closes these axioms, and the PLG axioms of identity and non-circularity, under the PLG rules (definition 25). It contains statements such as

$$0 = 0 < \mathsf{T}\ulcorner\mathsf{T}\ulcorner 0 = 0\urcorner\urcorner \tag{47}$$

and

$$4 + 3 \neq 5 \leqslant \mathsf{T}\dot{\neg}\dot{\mathsf{T}}\ulcorner 4 + 3 = 5\urcorner \tag{48}$$

**Proposition 14.** *GT is sound and complete with respect to the structure $\mathfrak{T}$ of lemma 12. For all $\mathcal{L}_{\mathrm{ta}}$-sentences $\phi, \psi$,*

$$\phi <_{T} \psi \textit{ iff } `\phi < \psi' \in GT \tag{49}$$
$$\phi \prec_{T,\wedge} \psi \textit{ iff } `\phi \prec \psi' \in GT \tag{50}$$
$$\phi \leqslant_{T} \psi \textit{ iff } `\phi \leqslant \psi' \in GT \tag{51}$$
$$\phi \preccurlyeq_{T,\wedge} \psi \textit{ iff } `\phi \preccurlyeq \psi' \in GT \tag{52}$$

*Proof.* We prove soundness as in the argument for lemma 12, noting that the T-axioms are true in $\mathfrak{T}$ by design and do not have false PLG-theorems.

For completeness, cases (51) and (52) are either trivial or follow from the first respectively second case by subsumption. To show (49) and (50), recall firstly that a **T**-priority tree of $\phi$ is just a sequence of $\mathcal{L}_{\mathrm{ta}}$-sentences $\psi_0, \dots, \psi_n$ such that $\psi_{i+1} = T\ulcorner\psi_i\urcorner$ or $\psi_{i+1} = \neg T\ulcorner\phi\urcorner$ and $\psi = \neg\psi_i$ (a '**T**-sequence from $\psi$ to $\phi$').

(49) Assume $\phi <_{\mathbf{T}} \psi$. Then there is a **T**-sequence from $\psi$ to $\phi$. We reason by induction on its length. If it is 2 then $\phi$ is **T**-generated from $\psi$. In this case, $\phi < \psi$ is an axiom of GT. If the **T**-sequence is of length $n+1$ then for some $\zeta$, $\phi\mathbf{T}\zeta$ and there is a **T**-sequence of length $n+1$ from $\zeta$ to $\psi$. By our induction hypothesis, GT therefore contains $\zeta < \psi$. Since $\phi\mathbf{T}\zeta$ ensures that we also have $\phi < \zeta$ in GT, its derived rule Cut(<) gives the desired $\phi < \psi$.

(50) Analogously, using the subsumption rule $S(\frac{\leqslant}{\preccurlyeq})$. $\square$

Thus, the formal connection between **T**-groundedness and the notion of grounding as regimented by the theory GT is that between a model and its complete theory. Fine's PLG is the most advanced regimentation of the in-virtue-of relation available. Thus, proposition 14 is a strong technical base for my proposal: to understand the philosophical content of the truth generator **T** in terms of the in-virtue-of relation.

## 8.3 LOGIC

In the previous section I made a case for understanding the Kripkean truth generator **T** in terms of the in-virtue-of relation from chapter 7. However, semantic groundedness is not a matter of the truth generator **T** alone, but arises from its interplay with the derivation of complex truths from true literals. I now turn to this second component of my *high-resolution* analysis of semantic groundedness, the logic generators (§ 3.4). In this section, I argue that certain logic generators are also well understood in terms of the in-virtue-of relation. I will identify their intimate connection to formal principles that govern the interaction of *in-virtue-of* and logic.

However, here the situation is more complicated than in the previous section. For one, the formal concept of semantic groundedness comes in many varieties, say Weak Kleene logic as well as Cantini supervaluation. For another, there is only little work on how the in-virtue-of relation interacts with logic, and even less of it arrives at definite verdicts.

I address the former complication by focusing on the Strong Kleene variant of semantic groundedness. There are at least two reasons to do so. Firstly, most authors, including Kripke, were interested primarily in the Strong Kleene variant. Secondly, my analysis of chapter 3 showed that sentences in the Strong Kleene least fixed point are grounded by the truth generator **T** in combination with the Tarski generator **W**. **W**, again, generates the complete classical theory of a model from the literals true in it (fact 1 on p. 29). To this extent, **W**-groundedness is groundedness by classical logic. Therefore, it is appropriate to focus on the logic generator **W**.

The second complication mentioned above, the absence of definite principles as to how the in-virtue-of relation interacts with logic, I will address by extracting, from the literature, what consensus there is. Before I do so, however, let me make one disclaimer. What follows touches on deep and rich matters at the core of metaphysics, philosophy of language and philosophy of logic. Of course, I will not do justice to all aspects relevant to it. My agenda is very specific, and I will focus on material that directly bears on my project. I can do so because I my goal is not a comprehensive study of how the in-virtue-of relation interacts with logic, but an outline of one way of enriching semantic groundedness by such considerations.

I focus on two recent pieces which discuss the formal interaction of the in-virtue-of relation and logic, Fine [2010] and Schnieder [2011] The latter I cited already in the previous section. Now, I turn to how Schnieder motivates his formal system for the in-virtue-of operator 'because'. He assumes that if a sentence is 'governed by a classical truth-functional connective' then it 'has its truth value *because of* the truth values of the embedded sentences' [Schnieder, 2011, p. 448, his emphasis]. This thought, which he labels the 'core intuition', thus provides a sufficient condition for a complex truth to hold in virtue of simpler ones. For example, it allows us to say that it is not the case that snow is not white in virtue of snow being white.

Two comments are in order. For one, Schnieder's principle only applies to truths whose main connective is truth functional. Further, it only concerns *classical* truths. For another, as Schnieder himself points out [2011, p. 449], the core intuition needs not to be stated in the formal mode. We need not speak of $\phi \wedge \psi$'s truth value, but may say simply that $\phi \wedge \psi$ because $\phi, \psi$.

Can I use Schnieder's *core intuition* to argue for the philosophical significance of the logic generator **W**? Unfortunately not. It overshoots and implies *in-virtue-of* claims that we would not want to make, and that at any rate are not reflected by the generator. To see this, firstly note that taken at face value, Schnieder's statement of the core intuition quoted above says that false constituent clauses matter just as much as truths. After all, he writes that a sentence '[...] has its truth value because of *the truth values* of *the embedded* sentences' [Schnieder, 2011, p. 448, my emphasis]. For example, the disjunction 'Snow is white or grass is blue' is true just as much in virtue of the falsity of 'Grass is blue' as in virtue of the truth of 'Snow is white'.[6] Now, ask how the falsity of embedded sentences ought to be put in the material mode. The natural way is in terms of negation. Thus, Schnieder's principle implies that snow is white or grass is blue because snow is white, as well as because grass is not blue. This, I claim, is not intuitive. We can argue against it along the lines of my case above for the *true because* claim (p. 134). The disjunction is fully explained by saying that snow is white, and adding that grass is not blue does not improve on this explanation.

At any rate, such consequences of Schnieder's principle stand in the way of how I propose to account for semantic groundedness. The logic generator of semantic groundedness **W** does not allow us to generate a disjunction from one disjunct together with the negation of the other, whereas Schnieder's *core intuition* suggest it to hold in virtue of both.[7]

In order to account for **W** as in the previous section I accounted for **T**, I need principles about logic and the in-virtue-of relation that are less inclusive than the *core intuition*.

The difficulty with Schnieder's core intuition is naturally answered by saying that not all sentences embedded in some sentence matter for the truth of it. Which do? At this juncture it is useful to turn to a discussion in Fine [2010, p. 105]. Fine argues that certain principles about logic and the in-virtue-of relation are well motivated from two theses.

These are, firstly, that every truth expressed by a syntactically complex sentence, holds in virtue of some truths, its grounds. Secondly, 'the classical truth-conditions should provide us with a guide to ground' [ibid.]. How do we understand this best? By 'classical truth-conditions' I take Fine to mean the standard biconditionals such as: 'Snow is white or grass is blue' is true iff 'Snow is white' is true or 'Grass is blue' is true. They explain the truth of a complex sentence in terms of the truth of its subclauses. If a subclause fails to be true, its falsity matters only to the extent that we need to, so to speak, look elsewhere for truth. Thus, of the two values, truth is privileged. This distinguishes

6 My point can also be made as follows: the core intuition suggests a *Weak Kleene* truth table for disjunction, whereas Fine's theses below suggest a *Strong Kleene* table. It may help to think in these terms.

7 However, Schnieder's core intuition may well be used to account for the significance of the Weak Kleene generator (§3.5), and its corresponding variant of semantic groundedness.

Fine's approach from Schnieder's *core intuition*, which suggested that false and true subclauses matter equally.

As a consequence, Fine's theses avoid the undesirable implications of the *core intuition*. For example, given the truth that snow is white or grass is blue, the first thesis requires merely some ground, while the *core intuition* requires that the disjunction is true because snow is white as well as grass is not blue. By the second thesis, it suffices to note that classically, a disjunction is true if one of its disjuncts is. Now, snow is white, and by Fine's theses we are entitled to say that snow is white or grass is blue in virtue of snow being white. Generally, a disjunction is true in virtue of its true disjuncts, and if merely one of them is true then it, and only it, is in virtue of what the disjunction holds. This thought is formalized well as the following schemata for a language $\mathcal{L}_g$ with in-virtue-of operator '$<$'.

$$\begin{aligned} &\phi < (\phi \vee \psi) && \text{if } \phi \text{ is true} \\ &\phi < (\psi \vee \phi) && \text{if } \phi \text{ is true} \end{aligned} \tag{53}$$

Only truths stand in the in-virtue-of relation. This prevents the principle from being stated in purely syntactic terms, but requires its restriction to true $\mathcal{L}_g$-sentences, a semantic condition. Of course, in order to further investigate the principle, this condition must be spelt out relative to an $\mathcal{L}_g$-model. Below, I will do just that for a model of particular interest to my study.

Turning to conjunctions, classically a conjunction is true if and only if so are both conjuncts. By the first thesis, however, this biconditional is given a 'direction' [Fine, 2010, p. 106], and read as saying that a conjunction holds in virtue of its conjuncts.[8]

$$\phi, \psi < (\phi \wedge \psi) \quad \text{if } \phi \text{ is true} \tag{54}$$

As to negation, its classical truth conditions say that a negation is true if and only if what is negated is false. This poses a problem if we wish to apply Fine's first thesis (see also [Fine, 2012b, p. 62]). Again, only truths stand in the in-virtue-of relation. Therefore, the right-hand side cannot be read as giving that in virtue of which the negation is true – after all, what is negated is said to be false.

Fortunately, however, the classical truth conditions of negation can be split into such principles whose right- as well as left-hand side ascribe truth, and thus can be read according to Fine's second thesis. I move directly to the formal $\mathcal{L}_g$-schemata which these principles motivate.

$$\begin{aligned} \phi &< \neg\neg\phi && \text{if } \phi \text{ is true} \\ \neg\phi, \neg\psi &< \neg(\phi \vee \psi) && \text{if } \neg\phi, \neg\psi \text{ are true} \end{aligned} \tag{55}$$

$$\begin{aligned} &\neg\phi < \neg(\phi \wedge \psi) && \text{if } \neg\phi, \neg\psi \text{ are true} \\ &\neg\phi < \neg(\psi \wedge \phi) && \text{if } \neg\phi, \neg\psi \text{ are true} \end{aligned} \tag{56}$$

8 Recall that the operator '$<$' is of variable arity (§ 25).

Together with the schemata from equations (53) and (54), these principles give an $\mathcal{L}_g$-variant of what Fine has called the *impure* logic of ground [Fine, 2012b, p. 58]. As such, they complement nicely the structural principles of Fine's *pure* logic (§7.3).

Interestingly, the rules that Schnieder (2011, p. 449) proposes for his operator 'because', and which he motivates from the core intuition, are well viewed as variants of (53) to (56). For example, he proposes a rule $\dfrac{\phi}{\neg\neg\phi \text{ because } \phi}$, which corresponds to the first schema from (55). Thus, Schnieder's formal system can be motivated from Fine's two theses. This suggests to read Schnieder's statement of the *core intuition* more charitably than I have done, and understand it as expressing a thought closely related to what I have extracted from Fine's two theses. At any rate, principles (53) to (56) are in line with what comes close to a consensus about how the in-virtue-of relation interacts with propositional logic.[9]

I postpone to look at schemata for quantified logic (p. 150 below), and instead now argue that the principles formalized in (53) to (56) have philosophical significance. Then, I will propose to view the logic generator **W** as exemplifying these natural principles, and thus being endowed with philosophical content itself. As in the previous section, I will support my proposal by a technical result.

What is expressed in (53) to (56) is philosophically significant for the same reason that the *true because* claim of the previous section is. It does philosophical work. Consider the principle that a disjunction holds in virtue of any true disjunct (the 'disjunction principle'). It is used by Kit Fine [2001] in the passage cited above on p. 128, where he presents ways of settling disagreement as to what holds in virtue of what. He points out that we have an intuitive access to such questions, and gives the disjunction principle as one piece of such knowledge.

Other authors make philosophical use of the disjunction principle, too. Shamik Dasgupta, in his unpublished yet widely noted [2013], uses the disjunction principle to establish certain key features of the in-virtue-of relation. For example, he argues against the view that if $A$ in virtue of it being the case that $B$ then necessarily, if $A$ then $B$.[10] His argument is from the disjunction principle, assuming that the fact that $C \vee D$ necessitates neither that $C$ nor that $D$.

Authors do not in the same manner use the principles concerning negation (equations 55 and 55). For some, this is due to their view that for it to be the case that $\neg\neg A$ just is for it to be the case that $A$ (Bolzano WL §209, Correia [2011]). Generally, I suspect, the absence

9 Similar schemata or rules are endorsed in [Batchelor, 2010, p. 69] and [Rosen, 2010, p. 117]. Very recently [2013] Fabrice Correia has identified rules for a relation of logical in-virtue-of that can also be brought into this form, as I will explain below.

10 However, he does endorse, as many authors do, that in such cases, the fact that $B$ necessitates that $A$: necessarily, if $B$ then $A$. Leuenberger [2014] challenges this received opinion.

of negation examples in the literature is due to that they are less easily parsed than, e.g., the disjunction principles. However, I do not take this to question the philosophical significance of the negation principles; in particular as they are well motivated from the same considerations as the others.

I now turn to the principle that a conjunction holds in virtue of its conjuncts (the 'conjunction principle'). At two occasions, Bolzano uses it as evidence in support of his theory of *Abfolge*. In WL §199, he argues that even if a truth follows from its grounds by some general principle of valid inference, this general principle is not itself one of the grounds. His example is that Socrates is Athenian and philosopher in virtue of Socrates being Athenian and Socrates being philosopher, but not in virtue of a conjunction being derivable from its conjuncts.[11]

Yet another case is found in WL §222, where Bolzano discusses a candidate sufficient condition for one truth holding in virtue of some others. Reconstructing his argument goes beyond the scope of the present study (but see Rumberg and Roski [2012]); my interest is in that his example of a clear case of *Abfolge*, indeed his only example at this point, is that of a conjunction holding in virtue of its conjuncts.

In contemporary literature, we find similar applications of this conjunction principle. For example, Roderick Batchelor [2010] uses it to point out that the in-virtue-of relation is stricter than mere logical implication: while the fact that A and B implies that A, by the conjunction principle and irreflexivity, the latter does not hold in virtue of the former.

My proposal is this. Just as the set generator **S** is well read as expressing the thought that a set is constituted from its elements, and just as the truth generator **T** expresses the *true because* claim, the logic generator **W** expresses the principles expressed in (53) to (56). This distinguishes **T**-**W**-groundedness from philosophically vacuous instances of the general concept, and answers the challenge of chapter 5.

As in the previous section, this simple thought will be buttressed by a technical result. I will take (53) to (56) as axiom schemata, and show that their theory is sound and complete with respect to the priority relations of semantic groundedness. For this, I firstly transfer the principles above into schemata for the formal language $\mathcal{L}_{\text{gta}}$. In fact, I only need to translate those principles which concern negation and disjunction, since I assume these to be the only sentential connections in $\mathcal{L}_{\text{gta}}$, and conjunction to be defined in terms of them. As in

11 For accuracy, it must be kept in mind that Bolzano did not have the concept of a truth-functional connective, nor thought that the logical form of a proposition should be given in something like a modern formal language. However, what he says about those truths concerning Socrates is an instance of what we nowadays express by a schema like (54).

the previous section (p. 137), I formalize the restriction to truths as restriction to sentences in the least fixed point.

$$\begin{aligned} \phi &< \neg\neg\phi && \text{if } \phi \in I_{\mathbf{T}\text{-}\mathbf{W}} \\ \neg\phi, \neg\psi &< \neg(\phi \vee \psi) && \text{if } \neg\phi, \neg\psi \in I_{\mathbf{T}\text{-}\mathbf{W}} \end{aligned} \tag{57}$$

$$\begin{aligned} \neg\phi &< \neg(\phi \wedge \psi) && \text{if } \neg\phi, \neg\psi \in I_{\mathbf{T}\text{-}\mathbf{W}} \\ \neg\phi &< \neg(\psi \wedge \phi) && \text{if } \neg\phi, \neg\psi \in I_{\mathbf{T}\text{-}\mathbf{W}} \end{aligned} \tag{58}$$

Transferring the disjunction principle (53) into the setting of semantic groundedness is slightly more involved. Recall the concept of $\beth$-$gg$-*rank* relative to some generator $\beth$ and things $gg$ (p. 26). In particular, every $\mathcal{L}_{ta}$-sentence $\mathbf{T}$-$\mathbf{W}$-grounded in the true base language literals $\Lambda$ has a unique $\mathbf{T}$-$\mathbf{W}$-$\Lambda$ rank corresponding to, intuitively, how long it takes to generate it from the $\Lambda$. For my results below to go through, I have to require that a disjunction holds in virtue of one of its true disjuncts only if this disjunct is of lower $\mathbf{T}$-$\mathbf{W}$-$\Lambda$ rank.

$$\begin{aligned} \phi &< (\phi \vee \psi) && \text{if } \phi \in I_{\mathbf{T}\text{-}\mathbf{W}} \text{ and } \phi \text{ has lower } \mathbf{T}\text{-}\mathbf{W}\text{-}\Lambda \text{ rank than } \phi \vee \psi \\ \phi &< (\psi \vee \phi) && \text{if } \phi \in I_{\mathbf{T}\text{-}\mathbf{W}} \text{ and } \phi \text{ has lower } \mathbf{T}\text{-}\mathbf{W}\text{-}\Lambda \text{ rank than } \phi \vee \psi \end{aligned} \tag{59}$$

Let the theory GTPL of the *Grounds of Truth and Propositional Logic*, be the least set of $\mathcal{L}_{gta}$-sentences containing every instance of schemata 57 and 59 (the 'logic axioms'), the axioms of Fine's *Pure Logic of Ground* PLG (p. 124), as well as the truth axioms from page 137:

$$\begin{aligned} \phi &< T\ulcorner\phi\urcorner && \text{if } \phi \in I_{\mathbf{T}\text{-}\mathbf{W}} \\ \neg\phi &< \neg T\ulcorner\phi\urcorner && \text{if } \phi \in I_{\mathbf{T}\text{-}\mathbf{W}} \end{aligned}$$

and closed under the PLG rules.[12]

In the previous section, I observed that the theory GT is both sound and complete with respect to the relations of $\mathbf{T}$-priority. Can we make an analogous observation for GTPL? This theory concerns the grounds of truth and propositional logic. Recall how truths of propositional logic are generated from true literals by the generator $\mathbf{V}$ (§ 2.3). For the time being, I will consider its combination with Kripke's truth generator $\mathbf{T}$, and connect it to GTPL, analogously to how in the previous section the pure logic of ground was connected to the truth generator $\mathbf{T}$.

My starting point is to observe that the logic axioms in (53) to (55) correspond precisely to rules in terms of which $\mathbf{V}$ was given (p. 28). To see this, it helps to consider how the logic axioms relate to the rules of logical grounding proposed very recently by Correia (2013).

12 Note that because the axiom schemata involve complicated semantic conditions, GTPL is not computably enumerable and thus a theory only in a relaxed sense.

Take any of the axioms from (57) to (59) and rewrite it in rule form. For example, rewrite the schema $\phi < (\psi \vee \phi)$ as the rule: $\dfrac{\phi}{(\psi \vee \phi)}$ Thus, you arrive at the *basic rules* of Correia's system [2013, p. 4].

$$\frac{\neg\phi \qquad \neg\psi}{\neg(\phi \vee \psi)} \qquad \frac{\psi}{(\phi \vee \psi)}$$
$$\frac{\phi}{(\phi \vee \psi)} \qquad \frac{\phi}{\neg\neg\phi} \tag{60}$$

Now, it suffices to recall equation 6 (p. 28) and note that Correia's *basic rules* characterize the generator of propositional truths **V**. Thus, from the logic axioms given above we arrive at the rules of the generator **V**.

We can also proceed in the other direction. Start with a **V** rule, say $\dfrac{\phi}{\phi \vee \psi}$, and note that it determines a principle of **V**-priority such as that $\phi$ is **V**-prior to $(\phi \vee \psi)$, in symbols:

$$\phi <_{\mathbf{V}} (\phi \vee \psi) \tag{61}$$

Here, replace '$<_{\mathbf{V}}$' by '$<$', and you obtain the second of the principles in (53). Similarly for the other **V** rules. Each captures a schema of **V**-generation isomorphic to one of Fine's principles, much like **T**-generation lines up with the truth axioms of the previous section. Thus, Fine's principles about the grounds of logically complex truths relate to Tarski truth generation **V** just like these truth axioms relate to Kripke truth generation **T**.

Does this observation suffice for the soundness of GTPL with respect to the **T**-**V** priority relations? The theory GTPL closes the logic axioms under the rules of the *pure* logic of ground. Does **T**-**V** priority satisfy these rules? An observation by Fine himself, transferred to the present context, shows that they do not [Fine, 2012b, p. 58].[13] In particular, they violate the following rule of *amalgamation*.

$$\frac{\zeta_0, \ldots, \zeta_n < \phi \qquad \xi_0, \ldots, \xi_m < \phi}{\zeta_0, \ldots, \zeta_n, \xi_0, \ldots, \xi_m < \phi} \tag{62}$$

Amalgamation is derived within PLG (def. 25) by a subtle combination of Cut and Reverse Subsumption [Fine, 2012a, p. 7]. The relation of strict **V**-grounding, however, does not satisfy this rule, therefore fails to model the pure logic of grounding in the same way as **T** does.

Here is how to produce counterexamples to amalgamation. $(\phi \vee \psi)$ is **V**-grounded in $\phi$, and **V**-grounded in $\psi$. Now, in order for amalgamation to be satisfied, we need $\phi, \psi <_{\mathbf{V}} (\phi \vee \psi)$. However, **V** does not allow us to generate the disjunction from both disjuncts. Therefore, it is not the case that $\phi, \psi <_{\mathbf{V}} (\phi \vee \psi)$. The theory GTPL is closed under the pure logic of ground, in particular therefore under

13 I thank Jon Litland for friendly guidance through this territory.

amalgamation. Hence, the structure of the **V**-priority relations does not satisfy GTPL.

Recall, from section 2.5, the concept of *left-closed* generators (def. 7). Let us now call a generator $\daleth$ *amalgamating* if for any $xx, zz, y$, we have that $xx <_\daleth y$ and $zz <_\daleth y$ only if $xx, zz <_\daleth y$. Generalizing the observation of the previous paragraph, we find that $\daleth$ is amalgamating only if it is left-closed. **V** is not left-closed: $\phi\mathbf{V}(\phi \vee \psi)$ and $\psi\mathbf{V}(\phi \vee \psi)$ but not $\phi, \psi\mathbf{V}(\phi \vee \psi)$. Consequently, $<_\mathbf{V}$ does not satisfy the principles of the pure logic of ground.

Having noted that principles much like my logic axioms do not ensure amalgamation, Fine considers their extension by the principle that a disjunction with two true disjuncts holds in virtue of both [2012b, p. 29]. GTPL closes the logic axioms under the pure logic of ground. Therefore, every instance of (62) can be derived from the original axioms by the PLG amalgamation rule. Nonetheless, Fine's remark proves useful for my project as it suggests one way of obtaining a model for GTPL from the **V** priority relations.

Following our recipe, the principle (62) correspond to the following candidate supplementation of the **V** rules.

$$\frac{\phi \qquad \psi}{(\phi \vee \psi)} \tag{63}$$

Combined with the **V**-rules it gives a logic generator $\mathbf{V}^+$ that is easily seen to be left-closed in the sense of definition 7 (p. 36). At this point, we can make use of the following lemma, and infer from it that $<_{\mathbf{V}^+}$ satisfies amalgamation.

**Lemma 13.** *Every left-closed generator is amalgamating.*

*Proof.* Let $\daleth$ be any left-closed generator. Based on proposition 1, I show that $xx <_\daleth y$ and $zz <_\daleth y$ only if $xx, zz <_\daleth y$, by nested induction on the height of $\daleth$-$zz$-priority trees. For simplicity, I assume that the relevant well-ordering $ww$ is set-sized, and will work with its standard ordinal representation.

For the rest of this proof, I will suppress mention of $\daleth$ where context fixes matters. I will write $h(xx, y)$ for the height of the shortest $xx$-priority tree of $y$. Without loss of generality, assume that $h(xx, y)$ is less than or equalt to $h(zz, y)$. We reason by induction on $h(zz, y)$.

At the base, we know that $zz\daleth y$ and $xx\daleth y$. Since $\daleth$ is left-closed, $zz, xx\daleth y$, hence $xx, zz < y$. For the induction step, let $h(zz, y) = \alpha + 1$ and assume that for all $uu$, if $h(uu, y)$ less than or equal to $\alpha$ then, if $xx < y$ and $uu < y$ then $xx, uu < y$. Let $xx < y$ and $zz < y$. Note that for some $uu$ the shortest $zz$-tree $\mathcal{T}$ of $y$ has a subtree that is a $uu$-tree of $y$ of height less than or equal to $\alpha$. By our induction hypothesis we therefore know that $y$ has a $\daleth$-$xx, uu$-tree $\mathcal{T}^\star$ .

Now consider those sequences $\langle y, \ldots, u \rangle$ among $\mathcal{T}^\star$ such that $u$ is one of $uu$. Each has a proper extension $\langle y, \ldots, u, z \rangle$ among $\mathcal{T}$. I claim that these sequences together with $\mathcal{T}^\star$ are a $\beth, xx, zz$-tree $\mathcal{T}^\dagger$ of $y$, and therefore $xx, zz <_\beth y$, as desired.

To prove my claim, by definition 3, it suffices to note that, firstly, $y$ is the root of $\mathcal{T}^\dagger$ because it is of $\mathcal{T}^\star$, and no length one sequence has been added. Secondly, each sequence $\langle y, \ldots, as, u\rangle$ has an initial segment among $\mathcal{T}^\dagger$ because $\langle y, \ldots, u\rangle$ and each of its initial segments is among $\mathcal{T}^\star$. Finally, for every sequence $\langle z, u, \ldots, y\rangle$, $u$ is $\beth$-generated partly from $z$. because each of these sequences is among $\mathcal{T}^\star$, in which $y$ has a $\beth$-$zz$-tree. □

Recall what it means for two generators to interfere (def. 9 on p. 36) and note that **T** and $\mathbf{V}^+$ do not interfere. Hence, their combination is left-closed (p. 37) and therefore amalgamating. More generally, the combination of Kripke's truth generator **T** and the left-closure $\mathbf{V}^+$ of **V** provides priority relations that satisfy the rules of the *pure logic of ground*.

On this basis, I can now show that GTPL is sound and complete with respect to the **T**-$\mathbf{V}^+$ priority relations. Towards this, I firstly note that a if a sentence is **T**-$\mathbf{V}^+$-generated, then it is so from finitely many sentences. Generally,

**Lemma 14.** *For every $\mathcal{L}_{\text{ta}}$-sentence $\phi$, if $\phi$ is **T**-$\mathbf{V}^+$ grounded in some sentences then these are finitely many.*

*Proof.* Assume that $\phi$ is **T**-$\mathbf{V}^+$-grounded in some sentences $\psi, \ldots$, and that this is witnessed by a **T**-$\mathbf{V}^+$-$\psi, \ldots$-priority tree $\mathcal{T}$. In order to show that $\psi, \ldots$ can be enumerated $\psi_0, \ldots, \psi_n$, for some natural number $n$, we reason by induction on the height of $\mathcal{T}$. At its base, $\phi$ is generated from $\psi \ldots$ either by **T** or $\mathbf{V}^+$. Either way, $\phi$ is generated from at most two sentences.

For the induction step, assume that $\phi$ is **T**-$\mathbf{V}^+$-generated from some sentences, which by our induction hypothesis we know to be $\zeta_0, \ldots, \zeta_m$ for some $m$, and that each $\zeta_i$ is generated from some sentences among $\psi, \ldots$. As in the base case, we show that each $\zeta_i$ is generated from at most two sentences. Hence, $\psi, \ldots$ are at most $2 \times m$ many sentences. □

The fact that tracing back **T**-$\mathbf{V}^+$ generation does not confront us with an infinity of sentences, renders it possible for the finite theory GTPL to prove all facts about **T**-$\mathbf{V}^+$-priority, just like the theory GT of the previous section was not merely sound, but also complete with respect to **T**-priority.

**Proposition 15.** *The theory GTPL is sound and complete with respect to the structure $\mathfrak{R} = (I_{T\text{-}V^+}, <_{T\text{-}V^+}, \leqslant_{T\text{-}V^+} \prec_{T\text{-}V^+,\wedge}, \preccurlyeq_{T\text{-}V^+,\wedge})$.*

*For all $\mathcal{L}_{\text{ta}}$-sentences $\phi, \psi$*

$$\phi \prec_{T\text{-}V^+,\wedge} \psi \text{ iff } `\phi \prec \psi' \in GTPL \tag{64}$$

$$\phi \preccurlyeq_{T\text{-}V^+,\wedge} \psi \text{ iff } `\phi \preccurlyeq \psi' \in GTPL \tag{65}$$

*For all $\mathcal{L}_{ta}$-sentences $\phi_0, \ldots, \phi_n, \psi$ for some natural number $n$*

$$\phi_0, \ldots, \phi_n <_{\mathbf{T}\text{-}\mathbf{V}^+} \psi \text{ iff } `\phi_0, \ldots, \phi_n < \psi' \in \mathit{GTPL} \tag{66}$$

$$\phi_0, \ldots, \phi_n \leqslant_{\mathbf{T}\text{-}\mathbf{V}^+} \psi \text{ iff } `\phi_0, \ldots, \phi_n \leqslant \psi' \in \mathit{GTPL} \tag{67}$$

*Proof.* In a nutshell, the proposition holds because the axioms of GTPL capture the generators **T** and **V**, and PLG derives statements that express, in $\mathfrak{R}$, mediate and partial **T**-**V**$^+$-generation.

For soundness (right-to-left), we again need to show firstly that the GTPL axioms are true in $\mathfrak{R}$, and secondly, that the PLG rules preserve truth in $\mathfrak{R}$. The PLG axioms of identity and non-circularity, and the truth axioms $\phi < T\ulcorner\phi\urcorner$ and $\neg\phi < \neg T\ulcorner\phi\urcorner$ hold for the same reason as they hold in the structure $\mathfrak{T}$ (lemma 14).

The logic axioms hold in $\mathfrak{R}$ because, by the recipe given above (p. 145), they correspond to the rules of $\mathbf{V}^+$. For example, $\phi < \neg\neg\phi$ holds because $\neg\neg\phi$ has a simple **T**-**V**$^+$ tree from $\phi$: $\begin{array}{c}\neg\neg\phi \\ | \\ \phi\end{array}$.

To see that the PLG transitivity rules preserve truth in $\mathfrak{R}$, it suffices to recall how the **T**-**V**$^+$ priority relations are defined. The subsumption rules $S(\frac{\leq}{\leqslant})$ and $S(\frac{\preccurlyeq}{\preccurlyeq})$ hold by the definition of $\leqslant_{\mathbf{T}\text{-}\mathbf{V}^+}$ respectively $\preccurlyeq_{\mathbf{T}\text{-}\mathbf{V}^+,\Lambda}$. Similarly, the soundness of $S(\frac{\leqslant}{\preccurlyeq})$ follows once we have shown that $S(\frac{\leq}{\prec})$ preserves truth in $\mathfrak{R}$, i.e. that if GTPL contains $\zeta, \ldots < \phi$ then $\zeta, \ldots \prec_{\mathbf{T}\text{-}\mathbf{V}^+,\Lambda} \phi$. This is shown by induction on the height of $\phi$'s **T**-**V**$^+$ tree which witnesses the antecedent. Firstly, though, note that by the condition on GTPL's axiom schemata, $\phi$ is **T**-**V**$^+$-grounded in $\Lambda$, hence (proposition 1) it has some **T**-**V**$^+$-$\Lambda$ priority tree whose height equals $\phi$'s rank. What needs to be shown is that this tree $\mathcal{T}$ goes through $\zeta, \ldots$.

At the base of the induction on the height of that tree which witnesses the antecedent, $\phi$ is **T**-**V**$^+$-generated from $\zeta, \ldots$. Now, the only subtle cases are those of a *disjunction* $\phi$, in all others we know that $\mathcal{T}$ goes through $\zeta, \ldots$ (for example, if $\phi$ is **T**-generated from $\zeta$, it can only be generated thus). Dealing with disjunctions, however, it pays off that GTPL has axioms $\zeta < (\zeta \vee \xi)$ and $\zeta < (\xi \vee \zeta)$ for all and only those $\zeta$ whose **T**-**W**-$\Lambda$ rank is less than that of the disjunction (schemata 53). We therefore know that $\phi$'s rank is higher than that of the disjunct which it is generated from. This in turn ensures that if generated step-by-step and starting from $\Lambda$, $\phi$ is generated from $\zeta$. Hence, the tree $\mathcal{T}$ witnessing $\phi$'s groundedness in $\Lambda$, goes through $\zeta$.

For the induction step, assume that in $\mathcal{T}$, $\phi$ is **T**-**V**$^+$-generated from some $\psi, \ldots$ which in turn are **T**-**V**$^+$-generated from $\zeta, \ldots$. By the induction hypothesis, there is a **T**-**V**$^+$-$\Lambda$ priority tree $\mathcal{T}'$ whose height equals $\psi$'s rank, and one of whose leaves is $\zeta$. Extend this tree $\mathcal{T}'$ by generating $\phi$ from its root $\psi$; the result witnesses $\zeta \prec_{\mathbf{T}-\mathbf{V}^+,\Lambda} \phi$, as desired.

Finally, to show that the cut rule satsfies truth in $\mathfrak{R}$, it is key to note that we now work with the *amalgamating* generator $\mathbf{V}^+$. For reverse subsumption, we make again use of the irreflexivity of $\prec_{\mathbf{T\text{-}V}^+-\wedge}$ together with the definition of $\leqslant_{\mathbf{T\text{-}V}^+}$.

For completeness (left-to-right), as in the proof of proposition 14, cases (67) and (65) are either trivial or follow from the other cases by subsumption. Let us look at (64) and (66) separately. Firstly, to show (64) assume that $\phi \prec_{\mathbf{T\text{-}V}^+,\wedge} \psi$. Hence, some $\mathbf{T\text{-}V}^+$,$\wedge$-priority tree $\mathcal{T}$ of $\psi$ contains a sequence $\langle\psi, \dots, \phi\rangle$. We reason by induction on the length of this sequence.

At its *base*, we know that $\phi$ is among some sentences from which $\psi$ is $\mathbf{T\text{-}V}^+$-generated. Then, $\phi, \dots < \psi$ is either an axiom of truth, or of logic. Either way, by the subsumption rule $S(\substack{\leq\\\prec})$ the theory GTPL contains $\phi \prec \psi$. For example, if $\psi$ is of the form $\neg(\zeta \wedge \xi)$ we know that GTPL contains $\neg\zeta, \neg\xi < \psi$. Since $\psi$ is $\mathbf{T\text{-}V}^+$-generated partly from $\phi$, we also know $\phi$ to be either $\neg\zeta$ or $\neg\xi$. Then, however, the PLG rule $S(\substack{\leq\\\prec})$ allows us to infer $\phi \prec \psi$, as desired.

For the induction *step*, let $\phi \prec_{\mathbf{T-V}^+,\wedge} \psi$ be witnessed by $\langle\psi, \dots, \chi, \phi\rangle$, and assume that $\chi \prec \psi$ $\in$GTPL. By the same argument as in the base case, we show that $\phi \prec \chi$ $\in$GTPL. Then, by subsumption $S(\substack{\leq\\\preccurlyeq})$ and the transitivity rule $T(\substack{\preccurlyeq\\\prec})$ we derive $\phi \prec \psi$.

Secondly, for (66) assume that $\phi_0, \dots, \phi_n <_{\mathbf{T\text{-}V}^+} \psi$. Then $\psi$ has a $\mathbf{T\text{-}V}^+$-$\phi_0, \dots, \phi_n$-priority tree $\mathcal{T}$. We show that GTPL contains $\phi_0, \dots, \phi_n <$ $\psi$ by induction on $\mathcal{T}$'s height (p. 31). At its base we reason as in the case of (64), except that now we do not even need the subsumption rule but have $\phi_0, \dots, \phi_n < \psi$ directly as an axioms of truth or logic.

For the induction step, assume that $\zeta_0, \dots, \zeta_m <_{\mathbf{T\text{-}V}^+} \psi$ and (†) each $\zeta_j$ is $\mathbf{T\text{-}V}^+$-generated from some $\phi^j_k, \dots, \phi^j_l$, $k \leqslant l \leqslant n$, as well as that (‡) every $\phi_i$, $i \leqslant n$, is among some sentences from which some $\zeta_j$ is thus generated. By our induction hypothesis, we know that GTPL contains $\zeta_0, \dots, \zeta_m < \psi$. Then, by subsumption $S(\substack{\leq\\\leqslant})$, $(\substack{\leqslant\\\leqslant})$ Cut and reverse subsumption we derive $\phi^0_k, \dots, \phi^0_l, \dots, \phi^m_k, \dots, \phi^m_l < \psi$ (compare also [Fine, 2012a, p. 6]). By (†), each $\phi^j_l$ is some $\phi_i$, $i$ less than or equal to $n$; by (‡), each $\phi_i$, $i$ less than or equal to $n$ is some $\phi^j_k$, $j$ less than or equal to $m$. Consequently, the theory GTPL contains the $\mathcal{L}_{gta}$-sentence $\phi_0, \dots, \phi_n < \psi$, as desired. □

I now turn to how the in-virtue of relation interacts with the quantifiers. Since I assume '∃' in the language $\mathcal{L}_{gta}$ to be defined in terms of '∀', it suffices to give principles for universal quantification. I follow the same strategy as above (p. 140) and read the classical truth conditions of a universal quantification as giving in virtue of what it is true, thus endowing it with a *direction*, to use Fine's figure of speech.

However, this strategy encounters some complication which we did not meet in the propositional case [Fine, 2012b, p. 60f]. A sentence of the form $\neg\forall x\phi(x)$ is true if and only if something does not satisfy $\phi(x)$. Thus, a truth of this form holds in virtue of a single witness truth $\neg\phi(\overline{o})$, for some object o of the domain. So far so good. A universal quantification $\forall x\phi(x)$, however, is true if and only if every object satisfies $\phi$. Following our strategy, we therefore would like to say that a universal quantification holds in virtue of all these truths.

To accommodate this thought, Fine offers two distinct axioms (ibid.). One is formulated in terms of a primitive relation "... are all and only the objects that exist". Then, $\forall x\phi(x)$ is grounded in $\phi(\overline{o}), \ldots, \phi(\overline{p})$ and the fact that $o, \ldots, p$ are all and only the objects that exist. The other, simpler rule requires a non-free background logic, and that the language provides a name for every object of the domain. These assumptions are reasonable for my purpose, as I want to apply the principles to the language $\mathcal{L}_{ta}$ of semantic groundedness. Here, every object of the intended domain has a name, since for every standard number $n$ its numeral $\overline{n}$ is an $\mathcal{L}_{ta}$-term.

The idea then is that from $\mathcal{L}_{ta}$-truths $\phi(\overline{0}), \phi(\overline{1}), \ldots$, we can infer that the truth $\forall x\phi$ holds in virtue of them. This thought is natural enough, but it cannot be implemented in the present setting. In $\mathcal{L}_{gta}$, the in-virtue-of relation is expressed by a sentential operator. Thus, implementation requires us to acknowledge $\mathcal{L}_{gta}$-sentences of infinite length, since $\forall x\phi$ holds in virtue of infinitely many truths; but of course $\mathcal{L}_{gta}$ is an ordinary finite language.

My present goal is to account for the philosophical significance of semantic groundedness and I use the in-virtue-of relation for this. However, I do not aim this account to be spelt out in full detail. In particular, as noted at the end of the previous chapter (p. 125), I do not attempt to formulate Fine's *pure logic of ground* in its intended infinitary setting. Now I have found that an infinitary language is needed in order to formalize adequately how, according to Fine, the in-virtue-of relation interacts with the universal quantifier. Accordingly, I will work with only a fragment of Fine's impure logic, given by the following axioms. I write $\phi(a)$ to indicate a sentence that contains the term $a$, and $\phi(x)$ to indicate a formula with a single free variable $x$.

$$\begin{array}{lr} \phi(a) \prec \forall x\phi(x) & \text{if } \phi(a), \forall x\phi(x) \text{ are true} \\ \neg\phi(a) < \neg\forall x\phi(x) & \text{if } \neg\phi(a) \text{ is true} \end{array} \tag{68}$$

Let GTL be the least set of $\mathcal{L}_{gta}$-sentences containing the PLG axioms, the truth axioms, the axioms governing the interaction of the in-virtue-of relation with propositional logic, as well as

$$\phi(a) \prec \forall x\phi(x) \qquad \text{if } \phi(a), \forall x\phi(x) \in I_{\mathbf{T\text{-}W}} \tag{69}$$

$$\neg\phi(a) < \neg\forall x\phi(x) \qquad \begin{array}{l} \text{if } \neg\phi(a) \in I_{\mathbf{T\text{-}W}} \text{ and has} \\ \quad \text{lower } \mathbf{T\text{-}W}\text{-}\Lambda \text{ rank than } \neg\forall x\phi(x) \end{array} \tag{70}$$

After having above shown the theory GTPL to be sound and complete with respect to the structure $\mathfrak{R}$ of $\mathbf{T\text{-}V}^{+}$ priority relations, I now turn to ask if GTL can be shown sound and complete with respect to $\mathbf{T\text{-}W}$ priority.

Firstly, let us ask for soundness. Of course, since **V** needed to be supplemented by the generator given in equation 63, so does **W**. However, more is needed. In the present, quantified setting, the problem concerning the amalgamation rule derived in PLG shows up in a second guise.

Consider the following two instances of principle (70): $1 \neq 1+1 < \neg\forall x(x = x+1)$ and $2 \neq 2+1 < \neg\forall x(x = x+1)$. By amalgamation, GTL contains $1 \neq 1+1, 2 \neq 2+1 < \neg\forall x(x = x+1)$. The Tarski truth generator **W**, however, does not allow us to generate the right-hand side from both sentences on the left-hand side. Therefore, it is *not* the case that $1 \neq 1+1, 2 \neq 2+1 <_{\mathbf{W}} \neg\forall x(x = x+1)$.

This analogous problem has a just analogous solution. We supplement **W** by the following way of generating negated universal quantifications [Fine, 2012b, p. 59].

$$\frac{\neg\varphi(\overline{n}) \qquad \ldots \qquad \neg\varphi(\overline{m})}{\neg\forall x\varphi}\; n, \ldots, m \text{ are some numbers} \tag{71}$$

Let $\mathbf{W}^+$ be the combination of **W** with the generators given in (63) and (71). I will show that the **T**-$\mathbf{W}^+$ priority relations satisfy the theory GTL. In fact, we can show that the connection between the principles that govern the in-virtue-of relation, and **T**-$\mathbf{W}^+$ groundedness, is even stronger. However, unlike in the case of GT and GTPL, GTL does not fully capture **T**-$\mathbf{W}^+$ priority. Propositions 14 and 15 are completeness results. For example, GT contains $\varphi < \psi$ exactly if $\psi$ is **T**-grounded in $\varphi$. Can we similarly prove completeness of GTL with respect to **T**-$\mathbf{W}^+$ priority?

Unfortunately not. A true universal quantification, such as $\forall x, x = x$ is **W**-grounded in infinitely many truths. Since our language of the in-virtue-of relation, however, is finite, the corresponding statement that $\forall x\, x = x$ is true in virtue of $0 = 0, 1 = 1, \ldots,$ cannot be expressed, much less proved.[14] GTL has axioms merely of what a universal quantification holds *partially* in virtue of, for instance $0 = 0 \prec \forall x\, x = x$. Accordingly, the best we can hope for is that GTL contains such statements $\varphi \prec \psi$ for every corresponding case of **T**-$\mathbf{W}^+$ dependence $\varphi \prec_{\mathbf{T}\text{-}\mathbf{W}^+,\wedge}$. And this can be proved.

**Proposition 16.** *The theory GTL is sound with respect to the structure* $\mathfrak{S} = (I_{\mathbf{T}\text{-}\mathbf{W}^+}, <_{\mathbf{T}\text{-}\mathbf{W}^+}, \prec_{\mathbf{T}\text{-}\mathbf{W}^+,\wedge})$*, and partly complete.*

$$\varphi, \ldots <_{\mathbf{T}\text{-}\mathbf{W}^+} \psi \qquad \text{if } `\varphi, \ldots < \psi\text{'} \in GTL \tag{72}$$

$$\varphi, \ldots <_{\mathbf{T}\text{-}\mathbf{W}^+} \psi \text{ or } \varphi, \ldots \text{ are } \psi \qquad \text{if } `\varphi_0, \ldots, \varphi_n \leqslant \psi\text{'} \in GTL \tag{73}$$

$$\varphi \prec_{\mathbf{T}\text{-}\mathbf{W}^+,\wedge} \psi \qquad \text{iff } `\varphi \prec \psi\text{'} \in GTL \tag{74}$$

$$\varphi \prec_{\mathbf{T}\text{-}\mathbf{W}^+,\wedge} \psi \text{ or } \varphi = \psi \qquad \text{iff } `\varphi \preccurlyeq \psi\text{'} \in GTL \tag{75}$$

Importantly, lines 72 and 73 are not biconditionals, but only state the soundness of GTL with respect to **T**-$\mathbf{W}^+$ grounding.

14 Also, no analogue of lemma 14 above is available.

*Proof.* Soundness is proved analogously to above. We note that the axioms (69) and (70) are true in $\mathfrak{S}$, because $\forall x\phi(x)$ is partly $\mathbf{W}^+$-generated from $\phi(\overline{n})$, and $\neg\forall x\phi(x)$ is fully $\mathbf{W}^+$-generated from $\neg\phi(\overline{n})$.

Completeness, too, is largely proved analogously to the proof of proposition 15. To show that '$\phi \prec \psi$'$\in$ GTL if $\psi$ $\mathbf{T}$-$\mathbf{W}^+$-depends on $\phi$, we reason by induction on the length of the witness sequence $\langle\psi,\ldots,\phi\rangle$. At its base, $\psi$ is $\mathbf{T}$-$\mathbf{W}^+$-generated from sentences $\phi,\ldots$. Either $\psi$ is a universal quantification or not. If not, then $\phi,\ldots < \psi$ is a GTL axiom, and we obtain $\phi \prec \psi$ by the PLG subsumption rule. If $\psi$ is a universal quantification, and $\phi$ is one of the infintely many sentences from which it is generated, then $\phi \prec \psi$ is a GTL axiom, and we are done.

For the induction step, let $\phi \prec_{\mathbf{T}\text{-}\mathbf{V}^+\text{-}\wedge} \psi$ be witnessed by a sequence $\langle\psi,\ldots,\chi,\phi\rangle$, and assume that '$\chi \prec \psi$'$\in$ GTL. By reasoning just analogously to the base, we derive $\phi \prec \chi$, and infer $\phi \prec \psi$ by the transitivity of $\prec$. $\square$

Proposition 16 shows that the structure of $\mathbf{T}$-$\mathbf{W}^+$-priority relations among the grounded truths mirrors their order by a relation of the kind as the theory GTL describes. Thus, proposition 16 provides mathematical support for my proposal that the generators $\mathbf{T}$ and $\mathbf{W}^+$ have a natural reading in those principles which philosophers have argued to govern how the in-virtue-of relation interacts with logic. Moreover, I have above presented cases in which these very principles are used by philosophers in their arguments. Thus, proposition 16 supports the connection between the formal concept of semantic groundedness, modulo its supplementation by the rules of (62) and (71), and philosophical work.

## 8.4 RELATED WORK

How does what I undertook in the previous sections relate to the recent work by Fabrice Correia and Kit Fine mentioned in the introduction [Fine, 2010; Correia, 2013]?

Firstly, I approach the technical material from the opposite direction. My starting point is the formal concept of semantic groundedness, and the need to account for its philosophical significance. Fine and Correia are both primarily interested in the in-virtue-of relation, and use Kripke's model constructions to test and clarify candidate principles governing this notion. Speaking figuratively, in Fine's and Correia's work rigour flows from the formal concept of groundedness to the philosophical notions of grounding. My goal however, is to have philosophical illumination flowing the other direction.

However, I believe that this difference between our projects is not due to disagreement on the subject matter, but merely a difference in perspective. The projects are compatible, and may indeed stimulate one another.

Secondly, unlike Correia [2013] I show that certain in-virtue-of principles are satisfied by the *high resolution* characterization of semantic groundedness. Correia works with what I called the *low resolution* characterization, that is its standard presentation in terms of a single jump operator (§3.3). Early in the present study (§3.4), I showed that such a Kripke jump is well analyzed by the combination of the truth generator **T** and a logic generator, e.g. **W**. On this high resolution we see that Kripke's semantic groundedness is given by rules that, as we saw in the previous sections, correspond to in-virtue-of principles. Correia endorses similar principles, but his low resolution approach to semantic groundedness does not provide him with corresponding rules. As a consequence, Correia is forced to say that literals true for interpreting 'T' by the jump of some set X hold in virtue of those literals true for interpreting 'T' by X [Correia, 2013, lemma 7.9]. For example, Correia must say that $T\ulcorner 0 \neq 1 \urcorner$ holds partly in virtue of $3 \neq 1 + 1$, because the latter is true for the empty interpretation of 'T' and the Kripke jump of the empty set contains $0 \neq 1$. This, however, conflicts with our antecedent grasp of the in-virtue-of relation. Working with the *high resolution* characterization I avoided such unintuitive commitments.

Thirdly, unlike Correia, I provide a model for the stricter, non-circular in-virtue-of relation as characterized by Fine's pure logic. The key to this result is my definition of **T**-**W**-$\Lambda$-dependence (def. 5 on page 33). In a nutshell, an $\mathcal{L}_{at}$ sentence $\phi$ depends on a sentence $\psi$ if and only if there is what I call a **T**-**W** priority tree and, importantly, this tree is as high as how long it takes to **T**-**W** generate $\phi$ from $\Lambda$. Only this latter condition rules out detours as depicted in figure 7 (p. 33). Thus, it ensures dependence to be irreflexive, a fact that has been crucial to the results of the previous sections.

Correia does not extract an irreflexive concept of dependence from Kripke's semantic groundedness. Consequently, the in-virtue-of relation which he interprets in Kripke's construction fails to be non-circular. Correia acknowledges this fact but takes it as evidence against the non-circularity of the in-virtue-of relation [Correia, 2013, p. 26]. I disagree, but postpone further discussion to another occasion, as presently my goal is just to explain how my approach differs from Correia's. Suffice it to say that I believe that without a condition such as I impose on **T**-**W** priority trees, Kripke's construction does not provide a model of the in-virtue-of relation. I find support for this view in the work of Kit Fine (2010). He also works with an irreflexive notion of semantic dependence, and arrives at it by similar means. Let me explain briefly.

Like Correia, Fine models principles of the in-virtue-of relation and truth in Kripke's least fixed point construction, but unlike Correia, not using its standard *low resolution* characterization. Instead, Fine works with what he calls its 'proof-theoretic' presentation. The sentences of the least fixed point are derived from the arithmetical truths by rules [Fine, 2010, pp. 110f]. For present purposes, these rules can be identified with those by which I gave the Tarski generator **W** (p. 28) and the Kripke generator **T** (p. 49). Thus, Fine works with a *high resolution* account of Kripke's construction similar to that of chapter 3 above.

On this basis, Fine models the notion of non-circular, mediate partial in-virtue-of in terms of *efficient* derivations [Fine, 2010, pp. 112]. As far as I can see, the relevant notion of efficiency is best rendered precise by the condition which I impose on **T**-**W** priority trees: a derivation tree of $\phi$ is efficient only if it is as high as how long it takes to **T**-**W** generate $\phi$ from $\Lambda$, i.e. only if its height equals $\phi$'s rank.[15] To this extent, the previous sections may well be viewed as elaborating on Fine's work.

Fourthly, I went beyond both Correia's and Fine's soundness results in that I also showed completeness. In section 8.2 I observed that the pure logic of ground together with the axiom schemata $\phi < T\ulcorner\phi\urcorner$ and $\neg\phi < \neg T\ulcorner\phi\urcorner$, is complete with respect to the **T** priority relations. Then, in section 8.3, I showed that if we also add analogous principles for propositional logic, the resulting theory is complete for **T**-**V** priority. Due to my focus on a finite language for the in-virtue-of relation, this completeness result does not fully carry over to quantified logic. Nonetheless, proposition 16 is stronger than what is noted by Correia or Fine.

Moreover, my completeness results are desirable not merely because they strengthen Fine's and Correia's findings. They also support my philosophical project and constitute a close formal connection between the in-virtue-of relation and Kripke's concept of semantic groundedness, closer than mere soundness.

Finally, unlike both Correia and Fine, I argued that in-virtue-of and groundedness are connected also through some informal principles which on the one hand underly the rules of the generators **T** and **W**, and on the other hand govern the interaction of truth respectively logic with the in-virtue-of relation. Thus, I supplemented the merely formal connection of soundness and completeness by what may be called an intensional connection between the in-virtue-of relation and semantic groundedness. It has not been identified by Fine, nor by Correia. In sum, the material of the previous sections reveals that the philosophical notion of one truth holding in virtue of others, and Kripke's formal work on truth, are more closely connected than what has yet been observed.

My specific interest mentioned above is therefore well met; I gave formal and philosophical evidence that semantic groundedness can be accounted for in terms of the in-virtue-of relation. More precisely, I made a case that the philosophical significance of semantic grounded-

15 My reading of Fine's is supported by his remarks in end note 13 [2010].

ness can be accounted for from the thought that it is true that snow is white in virtue of snow being white, paired with its analogues about in virtue of what a logically complex truth holds. Kripke's fixed point construction is significant because it tracks these principles, just as the cumulative hierarchy of sets receives significance from the thought that a set is constituted from its elements.

## 8.5 Conclusion

In this chapter, I made a simple proposal. Both truth and logic generator exemplify natural principles of the in-virtue-of relation. In section 8.2 I argued that the truth generator **T** is well read as expressing the *true because* claim. If using **T** we generate one truth from another, then this is read as saying that the former holds in virtue of the latter. Section 8.3 presented a case that the generator **W** analogously captures principles as to how the in-virtue-of relation interacts with logic. For example, the thought that it is not the case that snow is not white *because* snow is white gives philosophical content to **W**-generation of the former from the latter.

I do not claim this insight to be very original. Rather, I suspect such considerations to underlie what many people find attractive about Kripke's construction. My contribution is to have rendered it explicit. In addition, however, I supported this simple, natural reading of Kripke's construction by a series of technical results. They show that the structure of **T**-**W** priority relations among the grounded truths mirrors their order by the in-virtue-of relation, according to the natural principles about truth and logic.

To this extent, semantic groundedness receives philosophical significance much like, according to the iterative conception of chapter 6, set groundedness receives significance from the philosophical notion of constitution. In sum, to repeat my slogan from the introduction, the in-virtue-of relation is for semantic groundedness what constitution is for set groundedness. I have proposed an *iterative conception of truth*.

# 9 A MODAL LOGIC OF GROUNDEDNESS

## 9.1 INTRODUCTION

I propose using philosophical concepts of priority, in particular constitution (chapter 6) and the in-virtue-of relation (chapter 7), to supplement the formal concept of groundedness. More precisely, I use these philosophical notions to account for the philosophical significance of its paradigm instances: the well-foundedness of sets (§2.7) and Kripke's semantic groundedness (§3). However, this philosophical view of groundedness faces a challenge. In the present chapter, I will present the objection, clarify it and develop a response.

## 9.2 THE GHOST OF THE HIERARCHY

The challenge is closely linked to Kripke's 1975 remark that [Kripke 1975:714]

> [...] the induction defining the minimal fixed point is carried out in a set-theoretic meta-language, not in the object language itself. [...] The ghost of the Tarski hierarchy is still with us.

Accordingly, I will speak of the *ghost challenge*. It goes as follows. If we have to ascend to a meta-language in order to speak of groundedness (in particular, of its philosophical content), then the notion of grounded truth is not available to us in our own language.

This challenge must be distinguished from the problem known in the trade as *revenge* (see, e.g., Beall [2007]). The revenge objection to Kripke's theory of truth is to observe that it cannot express the fact that in the least fixed point model, the liar sentence is not true, or not determinately true. To begin with, there is textual evidence that Kripke saw revenge and ghost challenge as separate problems. This is the context of the quote given above (my emphasis).

> [...] the present approach certainly does not claim to give a universal language, and I doubt that such a goal can be achieved. Firstly, the induction defining the minimal fixed point is carried out in a set-theoretic meta-language, not in the object language itself. Secondly, there are assertions we can make about the object language which we cannot make in the object language. For example, Liar sentences are *not true* in the object language, in the sense that the inductive process never makes them true; but we are precluded from saying this in the object language by our interpretation of negation and the truth predicate.

Kripke gives two reasons to be skeptical whether a universal language can be achieved. The second is that in the object language, we cannot say that the liar sentence is not true – this, however, we know to be the case by looking at our model from the outside. This is the problem of revenge. The first reason given by Kripke, however, is that the process by which we have arrived at the model, and which captures the idea given informally by Kripke's story (see §3.2 above), cannot be carried out in the object theory. The idea, now, is the philosophical content of semantic groundedness, and the problem of expressing it in our own language is what I call the *ghost challenge*.

In addition to textual evidence, we can also argue directly that my ghost challenge is distinct from the revenge problem. Assume for the sake of the argument that we accept the revenge challenge. We accept that if we look at our theory from the outside, there are facts pertaining to truth which we cannot express using the truth predicate of our theory. Then, the challenge from groundedness being a meta-theoretic notion is still pressing. For, assume that it is true that in order to speak of semantic groundedness we have to ascend to a meta-theory, and that we cannot speak of the groundedness of truth in our own language. If so, groundedness would be of little relevance for logical-philosophical research into truth. Since, this project is ultimately not about truth in specific models of a certain object language, and by our assumption groundedness only applies to object language truth.

Of course I do not claim that any such "sandbox" investigation lacks relevance. It certainly is valuable, indeed indispensable. However, its results feed into philosophical research only to the extent that from our sandbox findings we can extrapolate to truth in our language. This step requires that the concepts at work can be transferred from the object language to our own language. The ghost challenge is that this cannot be done for groundedness.

In sum, while the revenge problem is about *how much* we can do with the grounded truth predicate, the ghost challenge is about *whether* the groundedness approach can be carried out in our own language in the first place. In response to it, the next section motivates a new way of expressing groundedness, one that is not meta-theoretic.

## 9.3 SIDESTEPPING THE GHOST

Recall that Kripke does not just give a definition. He also tells a story of how a speaker learns to use the truth predicate (§3.2). It is central to Kripke's picture that this learning process is presented as a process *over time*. This suggests a simple response to the ghost challenge. Although the formal definition of grounded truth is essentially meta-linguistic, we can at least express the intuitive component of semantic groundedness contained in Kripke's story: that the speaker understands $T\ulcorner\phi\urcorner$ only if she has understood $\phi$ earlier. Let us add tense operators to our language, and formalize this intuitive thought: $T\ulcorner\phi\urcorner \rightarrow$ *earlier*, $\phi$. Regimenting the temporal metaphor of Kripke's story also provides a way of expressing the philosophical account of grounded truth that I proposed in chapter 3. The order in which the speaker comes to understand sentences containing truth is the order of grounding. Below, I will develop a *tensed* theory of truth and show that it proves every instance of the schema $T\ulcorner\phi\urcorner \rightarrow$ *earlier* $\phi$ (corollary 5). To this extent, it formalizes the intuitive component of semantic groundedness.

Moreover, there is a strong case for this tensed schema that in the previous chapter I have labelled the *true because* claim (p. 134), and thus an argument that such a tensed theory of truth expresses the philosophical content of semantic groundedness developed in the previous chapter. After all, this account was developed in close analogy to the iterative conception of sets and the role of constitution in it. Now, in the philosophical literature on the iterative conception, constitution is frequently expressed in temporal terms.

At this point, the reader may sense a worry along the following lines. Truth is absolute. It does not make sense to speak of something *becoming* true. This worry results from taking the temporal vocabulary more literally than I intend it to be understood. Again, an analogy with the iterative conception of sets helps. As explained on p. 109, this vocabulary cannot be taken at face value. Nonetheless, it helps the philosopher to grasp the idea that the elements are prior to their set.[1] In much the same way, I propose to use tense to express the philosophical content of semantic groundedness. Tense is not used to express matter-of-fact temporal relations, but the priority of $\phi$ over $T\ulcorner\phi\urcorner$. Thus, my proposed reading of the tense operators not only suggests a response to the ghost challenge but in fact enables us to make philosophical use of the temporal metaphor in the first place.

As I draw the analogy to the iterative conception of sets, I ought to acknowledge that a more common characterization of it is not in modal terms, but either using Boolos-Shoenfield theory of stages [Boolos, 1971] or the Scott-Potter theory of levels [Potter, 2004]. Can we not, it may be asked, respond to the ghost challenge along analogous, extensional lines? I believe we cannot. A theory of stages, or levels, does not answer the ghost challenge in a satisfactory manner. However, I will return to the stage-theoretic contender in my discussion below (§9.7).

So far, I have outlined how I will express the intuitive component of semantic groundedness. In fact, however, the tensed approach taken below also provides a new characterization of Kripke's formal concept of groundedness. According to Kripke, the truth of a sentence is grounded if and only if it is in the least fixed point of his Strong Kleene jump operator. This least fixed point is approximated by a transfinite sequence of stages, which corresponds to the learning process mentioned earlier. Arguably, therefore, it is the stages that carry

1 Of course, the latinism 'prior' is itself temporal vocabulary.

the weight of Kripke's formal concept. As Albert Visser remarks in his insightful Handbook entry (1983, p. 180),

> [...] philosophically speaking the fixed point is not the *terminus ad quem* but the stages construction is the basic thing.

Below (9.6), I will show that all well-ordered models of my tensed theory of truth are isomorphic to the stages of Kripke's fixed point construction. In this precise sense, the theory captures also the formal concept of semantic groundedness.

In sum, I respond to the ghost challenge as follows: There is a way of expressing groundedness other than ascending to the meta-theory. We can use *tense* instead. English already has the vocabulary for this, and so would the language of our universal theory. But even if this was not the case, and we had to extend our language by temporal operators, this would still not mean to ascend to a meta-theory. Let me use the following metaphor. Whereas adding meta-theory is a *vertical* extension of our theory, my proposal is that of a *horizontal* enrichment. We do not need to let the ghost chase us up the hierarchy, we can sidestep it.

## 9.4 A LOGIC FOR GROUNDEDNESS

I now turn to implementing the proposal. I will formulate a tensed theory of grounded truth. However, what follows is merely one way of carrying out my idea, and I do not think that my philosophical proposal stands or falls by its success.

As in chapter 3, let $\mathcal{L}_{ta}$ be the language of first order arithmetic extended by a unary relation symbol 'T'. Recall that the language does not contain a primitive symbol '$\rightarrow$' and that we define the material conditional in terms of negation and disjunction.

I wish to enrich this familiar machinery by resources to express the groundedness of truth without ascending to a meta-theory. My goal is to enable a theory of grounded truth to express the priority of $\phi$ over $T\ulcorner\phi\urcorner$. I want it to be able to state that for $T\ulcorner\phi\urcorner$ to be true at some point, $\phi$ must have been true earlier. If so, I submit, we have found a way of expressing in the object language the *true because* claim, that, say, it is true that snow in virtue of snow being white.

In order to achieve this, I will *modalize* the first-order setting of truth. More precisely, I will add the resources of *tense logic*. Tense logic is formulated using two primitive modal operators, one operator looking *forwards* in time and reading "it will always be the case that ...", another operator looking *backwards* in time and reading "it has always been the case that ...".

Traditionally, the forwards looking operator is denoted '**G**'. However, since I use boldface capital letters already for the various generators of the formal concept of groundedness (chapter 2), my notation will deviate from the traditional presentation of tense logic. Semantically, the operator "it will always be the case that . . . " works precisely like the box operator in other modal logics. Therefore, it makes sense to use the standard symbol '$\square$' for it.

Its counterpart is the backwards-looking operator "it has always been the case that . . . ". Semantically, it works like the box operator, except that it works on the converse of accessibility. I will extend $\mathcal{L}_{ta}$ by two operator symbols $\blacksquare$ and $\square$ (for intuition: the language has a dark past but a bright future) to the language $\mathcal{L}_{tam}$ ('*m*' for *modalized*). As usual, we define $\blacklozenge\phi$ as $\neg \blacksquare \neg\phi$, and $\lozenge\phi$ analogously. Occasionally, I will use ◩ $\phi$ (read: "it is always the case that" $\phi$) as a meta-linguistic abbreviation for $\blacksquare\phi \wedge \phi \wedge \square\phi$, and ⬘ (read: 'it is sometimes the case that' $\phi$) as short for $\blacklozenge\phi \vee \phi \vee \lozenge\phi$ [Garson, 1984, p. 292]. Although it comprises two modal operators $\square$ and $\blacksquare$, the language $\mathcal{L}_{tam}$ is interpreted in ordinary models $(W, R, D, d)$ of first-order modal logic. $\square\phi$, on the one hand, holds at some point $w$ from $W$ iff it holds at every point R-accessible from $w$, that is, at every point $v$ such that $wRv$. $\blacksquare\phi$, on the other hand, holds at $w$ iff it holds at every point *conversely* accessible, that is, at every point $v$ such that $vRw$. In effect, $\square$ "looks forwards" and $\blacksquare$ "looks backwards".

This already allows us to express the first component of my intuitive gloss on groundedness. The truth of $T\ulcorner\phi\urcorner$ presupposes the truth of $\phi$, that is: $T\ulcorner\phi\urcorner$ only if $\phi$ earlier, that is: $T\ulcorner\phi\urcorner \rightarrow \blacklozenge\phi$. I will assume that time is well-ordered. Formally, I will restrict attention to models $(W, R, D, d)$ such that R well-orders $W$. Quantified modal logic is hard, both technically and philosophically. Fortunately, I do not have to deal with its subtleties. All I want to say is that *truth* changes over time. What our first-order quantifiers range over remains the same, and so does what our terms denote. Therefore, I can let the quantifiers be governed by standard, non-free first-order logic, and assume all terms to be *rigid designators* [Garson, 1984]. The result is a basic first-order logic of well-ordered time: *woq*. I recall some basic definitions.

**Definition 26** (Validity and Consequence in Quantified Modal Logic)**.** Let $\mathfrak{F}$ be any frame $(W, R)$ and $\mathfrak{M}$ be any model of first-order modal logic based on $\mathfrak{F}$, we say that a sentence $\phi$ is *valid* in $\mathfrak{M}$ iff for every $w \in W$, $\mathfrak{M} \models \phi[w]$. We call $\phi$ *valid* in $\mathfrak{F}$ (in symbols: '$\models_{\mathfrak{f}} \phi$') iff for every model $\mathfrak{M}$, $\phi$ is valid in $\mathfrak{M}$. Finally, let $\mathfrak{f}$ be a class of frames (e.g. the well-ordered frames) we say that $\phi$ is a *consequence over* $\mathfrak{f}$ of some set of sentences $\Sigma$ (in symbols: $\Sigma \models_{\mathfrak{f}} \phi$) iff for every model $\mathfrak{M}$ based on $\mathfrak{F} = (W, R)$ and every $w \in W$, if $\mathfrak{M} \models \Sigma[w]$ then $\mathfrak{M} \models \phi[w]$.

Thus, I will write $\Sigma \models_{woq} \phi$ if for every *woq*-model $(W, R, D, d)$ and every $w \in W$, if $(W, R, D, d) \models \Sigma$ then $(W, R, D, d) \models \phi$. Note that *woq* validates the Barcan formulae for both operators.

The first-order logic of well-ordered time is a powerful tool. For two structures $S$ and $S'$. I write $S \simeq S'$ for the statement that there is an isomorphism between S and $S'$, and $\mathfrak{N}$ for the standard model of arithmetic.

**Theorem 1** (Scott, Garson). *Let $\mathcal{L}_{\mathrm{am}}$ be the language of first-order arithmetic plus tense operators '$\square$' and '$\blacksquare$'. Extend $\mathcal{L}_{\mathrm{am}}$ by a unary predicate symbol 'N' to the language $\mathcal{L}_{\mathrm{amn}}$. There are $\mathcal{L}_{\mathrm{amn}}$ sentences $\Sigma$ such that the following holds.*

*Let $(W, R, D, d)$ a constant-domain model of the first-order logic of well-ordered time* woq. *Then we have that for every point $w \in W$,*

$$(W, R, D, d) \models \Sigma[w] \Rightarrow (D, d) \simeq \mathfrak{N}$$

*That is, every point that satisfies $\Sigma$ is in fact a* standard *model of arithmetic.*

*Proof.* See Garson (1984), sections 3.2.2 and 3.2.3.

(Sketch) Let $\Sigma$ comprise the following

N1 $\forall x \blacklozenge (Nx \wedge \blacksquare \neg Nx \wedge \square \neg Nx)$

- "Every object of the domain has the property N at exactly one time"

N2 $\boxbslash\, \forall x \forall y \big((Nx \wedge Ny) \to x = y\big)$

- "No two things have the property N at the same time"

Thus, every model of N1 and N2 will have an injective function from the (possibly non-standard) domain of $\mathfrak{M}$ into the well-ordered set $W$.

Note that for every model $\mathfrak{M} = (W, R, D, d)$ such that P1$\wedge$P2 is true at some $w \in W$, we have that $\mathfrak{M} \models \blacklozenge(N\overline{n} \wedge \diamond N\overline{m})[v]$, $v \in W$, just in case $N\overline{n}$ is true at an R-earlier point than $N\overline{n}$. Thus, N1 and N2 have allowed us to define, by $\blacklozenge(Nx \wedge \diamond Ny)$, a restriction $R^N$ of the well-ordered relation R to those points that some object of the domain is mapped to. Let us write $x\tilde{R}y$ as a meta-linguistic abbreviation of $\blacklozenge(Nx \wedge \diamond Ny)$.

$\Sigma$ contains another axiom.

N3 $\forall x \forall y \Big(y = S(x) \leftrightarrow x\tilde{R}y \wedge \forall z\big((z \neq y \wedge x\tilde{R}z) \to y\tilde{R}z\big)\Big)$

- "$y$ is the successor of $x$ just in case $y$ is the least $z$ $R^N$-greater than $x$"

Finally, add to $\Sigma$ the axioms of Robinson arithmetic, in particular:

Q0 $\forall x\big(S(x)x \neq 0\big)$

Q1 $\forall x\Big(x = 0 \vee \exists y\big(x = S(y)\big)\Big)$

Let $\mathfrak{M}$ be any model, $w \in W$ and assume that $(W, R, D, d) \models \Sigma[w]$. As said before, the fact that $(W, R, D, d)$ satisfies N1 and N2 implies that the objects of the domain D are mapped injectively to points

| | | | |
|---|---|---|---|
| | $\phi \to \Box \blacklozenge \phi$ | $\phi \to \blacksquare \Diamond \phi$ | |
| K$\Box$ | $\Box(\phi \to \phi) \to (\Box\phi \to \Box\phi)$ | $\blacksquare(\phi \to \phi) \to (\blacksquare\phi \to \blacksquare\phi)$ | K$\blacksquare$ |
| 4$\Box$ | $\Box\phi \to \Box\Box\phi$ | $\blacksquare\phi \to \blacksquare\blacksquare\phi$ | 4$\blacksquare$ |
| .3$\Diamond$ | $\Diamond\phi \wedge \Diamond\psi \to \Diamond(\phi \wedge \Diamond\psi) \vee \Diamond(\Diamond\phi \wedge \psi) \vee \Diamond(\phi \wedge \psi)$ | | |
| .3$\blacklozenge$ | $\blacklozenge\phi \wedge \blacklozenge\psi \to \blacklozenge(\phi \wedge \blacklozenge\psi) \vee \blacklozenge(\blacklozenge\phi \wedge \psi) \vee \blacklozenge(\phi \wedge \psi)$ | | |

Table 2: Axioms for linear time

in $W$. These are well-ordered by R, and $x\tilde{R}y$ expresses precisely this restriction $R^N$ of R. N3 ensures that $\overline{n}$ is mapped to the $R^N$-successor of what is mapped to $S(\overline{m})$.

Thus, Q0 ensures that '$\overline{0}$' denotes the $R^N$-least point $w_0$ in $W$, and Q1 that $R^N$ is an $\omega$-sequence. Hence, the domain at $w$ is (isomorphic to) the standard numbers. However, since we have to do with constant-domain models of tense logic, we in fact know that $(D, d) \simeq \mathfrak{N}$.

□

**Corollary 1.** *The logic $\vDash_{woq}$ is not axiomatizable.*

*Proof.* (Idea, for details see Garson (1984:294).) Analogously to how we infer the incompleteness of second-order logic from Gödel's incompleteness theorem and the fact that second-order arithmetic is categorical.

The key observation is that for the set of sentences $\Sigma$ from proposition 1 and any sentence $\phi$ of first-order arithmetic,

$$\vDash_{woq} \Sigma \to \phi \text{ iff } \mathfrak{N} \vDash \phi \tag{76}$$

for the standard model of arithmetic $\mathfrak{N}$. □

Hence, there is no proof procedure complete with respect to the first-order logic of well-ordered time *woq*. This stands in contrast with many other first-order logics of time, that have such complete axiomatizations. For example, the logic of linear time is complete. In the following, I will therefore develop an axiomatic theory of grounded truth on the basis of linear time. As we will see, it will allow us to approximate syntactically what we have just found to be beyond the reach of axiomatization, the logic of well-ordered time.

Unlike *woq*, this logic of linear time *lq* is axiomatizable. There are complete proof procedures for *lq*, for example systems of labelled tableaux [Priest, 2008, 14.7.12]. However, as I will largely reason about rather than within *lq*, it will prove useful to work with a Hilbert-style axiomatic proof system. Its axioms for linear time are presented in table 2. Thus, let $\Sigma \vdash_{lq} \phi$ if there is a proof of $\phi$ from $\Sigma$ in a Hilbert-style axiomatization of the first-order logic of linear time *lq*.

Since, of course, every well-ordering is a linear order but not vice versa, the logic of well-ordered time *woq* is strictly stronger than *lq*.

In particular, complete proof systems for *lq* are sound with respect to the logic of well-ordered time *woq*.

## 9.5 A MODAL LOGIC OF GROUNDED TRUTH

I now formulate a theory of truth MGT in the logic of linear time *lq*. My goal is to capture the notion of grounded truth over first-order arithmetic. Accordingly, MGT is based on first-order arithmetic. More precisely, it includes first order Peano Arithmetic ('PA'), whose induction schema we generalize to the extended language $\mathcal{L}_{\text{tam}}$.[2] Further, our modal logic is intended to express the step-by-step construction of a type-free truth predicate over arithmetic. The base theory, however, is not subject to this development. It holds at every stage of the construction, hence *necessarily* in our chosen modal logic. Consequently, I put a ' $\boxbslash$ ' ("always") in front of every PA axiom. Let 'APA' denote the resulting $\mathcal{L}_{\text{tam}}$ theory. As a result, MGT proves $\boxbslash\, \phi$ for every PA theorem $\phi$ in the language of arithmetic.

**Lemma 15.** *Let $\phi$ be any $\mathcal{L}_a$-sentence.*

$$\text{PA} \vdash \phi \Rightarrow \text{APA} \vdash_{\text{lq}} \boxbslash\, \phi$$

Now, I add axioms that govern how 'T' interacts with the modal vocabulary.

(Ground) $\Diamond(\forall x \neg Tx)$

APA and Ground couched in a first-order logic of linear time provide a general framework for theory of grounded truth. By itself, however, it leaves open whether anything at all becomes true at some point. What needs to be added now are axioms of truth-introduction.

My goal in the present chapter is specific. I aim for a theory that expresses semantic groundedness, in particular **JSK**-groundedness.[3] Thus, I do not want it just to say that more and more sentences become true, but that this happens according to the *Strong Kleene* jump. Hence, I need axioms that say how truth is introduced according this evaluation scheme. However, the logic of tense I have chosen as my framework is based on classical first-order logic. And it is generally desirable to remain within the classical setting. In sum, we need axioms that express truth introduction according to the Strong Kleene jump operator, and do so in classical logic.

Fortunately, such axioms are available, in the system KF (for *Kripke-Feferman*). They express, in the object language of arithmetic plus

2 Arithmetic is a convenient base. However, this choice of a base theory is not essential to what follows.

3 As in chapter 4, I work with the received *low resolution* characterization of semantic groundedness (§ 3.3) to render my proposal more accessible. In this connection, recall my remarks on p. 50.

'T', that T is closed under the semantic clauses of a partial model $\mathfrak{N}(X^+, X^-)$ [Halbach, 2011, p. 204]. As a consequence, KF characterizes precisely the fixed points of the Strong Kleene jump $\mathcal{J}_{sk}$ [ibid., theorem 15.15]. And, KF is a classical theory. So, I will use the KF truth axioms to characterize, in the object-language, how in the course of Kripke's model construction, more and more sentences become true.

For this, however, the KF axioms need to be modified in an important way. The reason is that, as they stand, they describe an arbitrary fixed point and not the step-by-step *construction*. For example, one KF axiom is that a sentence $\phi$ is true if and only if it is true that $\phi$ is true. What we would like to say, however is that if $\phi$ is true then it *will be* true that it is true that $\phi$ is true, and it is true that $\phi$ is true only if $\phi$ has been true *earlier*.

Precisely for this purpose, however, we have availed ourselves of tense logic. So, I will reformulate the KF axioms using its operators.[4]

TKF1 $\boxbslash\ \forall x \forall y\, ((Tx \underset{\cdot}{=} y \to \blacklozenge x = y) \wedge (x = y \to \Diamond Tx \underset{\cdot}{=} y \wedge \Box Tx \underset{\cdot}{=} y))$

TKF2 $\boxbslash\ \forall x \forall y\, \big((Tx \underset{\cdot}{\neq} y \to \blacklozenge x \neq y) \wedge \big(x \neq y \to \Diamond Tx \underset{\cdot}{\neq} y \wedge \Box Tx \underset{\cdot}{\neq} y)\big)$

TKF12 $\boxbslash\ \forall x \Big(\big(T\underset{\cdot}{T}x \to \blacklozenge(Tx \wedge \blacksquare\neg Tx)\big) \wedge \big(Tx \to \Diamond T\underset{\cdot}{T}x \wedge \Box T\underset{\cdot}{T}x\big)\Big)$

TKF13 $\boxbslash\ \forall x \Big(\big(T\underset{\cdot}{\neg}\underset{\cdot}{T}x \to (\blacklozenge T\underset{\cdot}{\neg}x \vee \neg \underset{\cdot}{Sent}_{ta}(x))\big) \wedge \big((T\underset{\cdot}{\neg}x \vee \neg \underset{\cdot}{Sent}_{ta}(x)) \to \Diamond T\underset{\cdot}{\neg}\underset{\cdot}{T}x \wedge \Box T\underset{\cdot}{\neg}\underset{\cdot}{T}x)\Big)$

Each axiom has the form of a universally quantified conjunction, the first of which concerns the past while the second concerns the future. The reader may wonder why the second conjunct, unlike the first, contains not just a diamond but also a box. The reason is that the underlying tense logic *lq* does not prove the schema of seriality $\Box\phi \to \Diamond\phi$.

Note the first conjunct of TKF12. It says that if it is true that it is true that $\phi$, not only at some point in the past it was the case that it is true that $\phi$, but in fact there was an earliest such point.

What about the connectives and quantifiers? In Kripke's construction, at every stage, truth is closed under Strong Kleene logic. This closure is expressed by the remaining KF axioms which govern the interaction of 'T' with the quantifiers and connectives other than '$\neg$'. Hence, I add these axioms as they are, merely putting an 'always' ('$\boxbslash$') in front.

TKF3 $\boxbslash\ \forall x\, (\underset{\cdot}{Sent}_{ta}(x) \to (T\underset{\cdot}{\neg}\underset{\cdot}{\neg}x \leftrightarrow Tx))$

TKF4 $\boxbslash\ \forall x \forall y\, (\underset{\cdot}{Sent}_{ta}(x \underset{\cdot}{\wedge} y) \to (Tx \underset{\cdot}{\wedge} y \leftrightarrow Tx \wedge Ty))$

4 '$\underset{\cdot}{Sent}_{ta}$' arithmetizes the syntactic property of being an $\mathcal{L}_{ta}$-sentence. Note that this property is definable in arithmetic by a formula all of whose quantifiers are bounded. Hence, PA$\vdash \underset{\cdot}{Sent}_{ta}(\ulcorner\phi\urcorner)$ iff $\phi$ is an $\mathcal{L}_{ta}$-sentence.

TKF5 ◩ $\forall x \forall y\, (\mathit{Sent}_{ta}(x \underset{\cdot}{\wedge} y) \to (T\underset{\cdot}{\neg}\, (x \underset{\cdot}{\wedge} y) \leftrightarrow T\underset{\cdot}{\neg} x \vee T\underset{\cdot}{\neg} y))$

TKF6 ◩ $\forall x \forall y\, (\mathit{Sent}_{ta}(x \underset{\cdot}{\vee} y) \to (Tx \underset{\cdot}{\vee} y \leftrightarrow Tx \vee Ty))$

TKF7 ◩ $\forall x \forall y\, (\mathit{Sent}_{ta}(x \underset{\cdot}{\vee} y) \to (T\underset{\cdot}{\neg}\, (x \underset{\cdot}{\vee} y) \leftrightarrow T\underset{\cdot}{\neg} x \wedge T\underset{\cdot}{\neg} y))$

TKF8 ◩ $\forall y \forall x\, (\mathit{Sent}_{ta}(\underset{\cdot}{\forall} y x) \to (T\underset{\cdot}{\forall} y x \leftrightarrow \forall z(\mathit{ClTm}(z) \to Tx(z/y))))$

TKF9 ◩ $\forall y \forall x\, (\mathit{Sent}_{ta}(\underset{\cdot}{\forall} y x) \to (T\underset{\cdot}{\neg}\underset{\cdot}{\forall} y x \leftrightarrow \exists z(\mathit{ClTm}(z) \wedge T\underset{\cdot}{\neg} x(z/y))))$

TKF10 ◩ $\forall y \forall x\, (\mathit{Sent}_{ta}(\underset{\cdot}{\exists} y x) \to (T\underset{\cdot}{\exists} y x \leftrightarrow \exists z(\mathit{ClTm}(z) \wedge Tx(z/y))))$

TKF11 ◩ $\forall y \forall x\, (\mathit{Sent}_{ta}(\underset{\cdot}{\exists} y x) \to (T\underset{\cdot}{\neg}\underset{\cdot}{\exists} y x \leftrightarrow \forall z(\mathit{ClTm}(z) \to T\underset{\cdot}{\neg} x(z/y))))$

The result is my *modal logic of grounded truth* MGT: always arithmetic, the axiom of ground, and KF turned into axioms of step-by-step truth introduction.

Note that I understand KF as not including the consistency of truth

$$\forall x(\mathit{Sent}_{ta}(x) \to \neg(Tx \wedge T\underset{\cdot}{\neg} x))$$

Often, this formula is added to the KF axioms, since it ensures a number of pleasing results. I do not have to do so though, since in the present, tensed context, consistency can be shown to follow. See proposition 18 below.

Note further that the first conjuncts of axioms TKF1, -2, -12 and -13 allow us, if we have established a certain atomic sentence, to introduce respectively iterate the truth predicate *later*. For example, TKF1 allows us to infer, from $x = y$, that at some point later in time, $Tx \underset{\cdot}{=} y$. More generally, we can show that whenever a sentence is ascribed truth at some point, then it will remain true thereafter.

**Lemma 16.**

$$\mathrm{MGT} \vdash_{lq} ◩\ \forall x\big(\mathrm{Sent}_{ta}(x) \to (Tx \to \Box Tx)\big)$$

*Proof.* (*sketch*) Induction on positive complexity (see p. 82) within MGT. Since the PA induction has been extended to the language with truth predicate, firstly the positive complexity of an $\mathcal{L}_{ta}$ formula $\phi$ is represented in PA by a functional term $PC^\bullet$ such that PA $\vdash PC^\bullet(\ulcorner\phi\urcorner) = \overline{n}$ iff the positive complexity of $\phi$ is $n$. Secondly, PA proves the following induction principle.[5]

$$\begin{aligned} &\forall x\Big(\mathit{Sent}_{ta}(x) \wedge \forall y\big(\mathit{Sent}_{ta}(y) \wedge PC^\bullet(y) \leqslant PC^\bullet(x) \wedge (Ty \to \Box Ty)\big) \\ &\qquad \to (Tx \to \Box Tx)\Big) \\ &\to \forall x\big(\mathit{Sent}_{ta}(x) \to (Tx \to \Box Tx)\big) \end{aligned} \tag{77}$$

5 See also [Halbach, 1996, pp. 40f].

The base cases are then taken care of by the axioms TKF1,-2,-12 and -13; the induction step by the compositionality axioms TKF4-11 and the induction hypothesis. □

**Corollary 2.**

$$\text{MGT} \vdash_{\text{lq}} \text{◩}\, \forall x \big(\text{Sent}_{\text{ta}}(x) \rightarrow (\neg Tx \rightarrow \blacksquare \neg Tx)\big)$$

As I will show in the next section, the theory MGT has natural models. They provide a strong case for it as a theory of *grounded* truth. However, already from a proof-theoretic point of view, MGT has several desirable properties, as we will see in the remainder of this section.

The axioms TKF1 through TKF13 are well viewed as a modalization of KF. It is natural to ask how the system MGT relates to the standard, non-modal theory KF. To answer this question, we translate the language of truth $\mathcal{L}_{\text{ta}}$ into the language $\mathcal{L}_{\text{tam}}$.

**Definition 27.** We define a mapping $(\cdot)^\star$ from the $\mathcal{L}_{\text{ta}}$-formulae into the $\mathcal{L}_{\text{tam}}$-formulae. Let $a, b$ $\mathcal{L}_{\text{ta}}$ be terms and $\phi, \psi$ be $\mathcal{L}_{\text{ta}}$ formulae.

$$\begin{aligned} (a = b)^\star &= a = b \\ (Ta)^\star &= \text{⬙}\, Ta \\ (\neg\phi)^\star &= \neg(\phi)^\star \\ (\phi \vee \psi)^\star &= (\phi)^\star \vee (\psi)^\star \\ (\forall x \phi)^\star &= \forall x (\phi)^\star \end{aligned}$$

For sets of $\mathcal{L}_{\text{ta}}$-sentences $\Gamma$ I write '$(\Gamma)^\star$' for the set of translations of each sentence in $\Gamma$.

I will prove the following proposition.

**Proposition 17.** *MGT interprets KF: if* KF $\vdash \phi$ *then* MGT $\vdash_{\text{lq}} (\phi)^\star$

In order to do so, two lemmata are needed. The first one says that the translation $(\cdot)^\star$ preserves provability.

**Lemma 17.** *For every set of $\mathcal{L}_{\text{ta}}$-formulae $\Gamma$ and every $\mathcal{L}_{\text{ta}}$-formula $\phi$, if there is a proof of $\phi$ from $\Gamma$ in first order logic then there is a proof of $(\phi)^\star$ from $(\Gamma)^\star$ in* lq.

$$\Gamma \vdash \phi \Rightarrow (\Gamma)^\star \vdash_{\text{lq}} (\phi)^\star$$

*Proof.* By an induction on the length of proof $l$, exploiting the fact that our mapping $(\phi)^\star$ translates connectives and quantifiers homophonically.

If $l = 1$, then $\phi \in \Gamma$ or $\phi$ is an axiom of first order logic. If $\phi \in \Gamma$, then we also have: $(\phi)^\star \in (\Gamma)^\star$, hence $(\Gamma)^\star \vdash_{woq} (\phi)^\star$. If $\phi$ is a logical axiom, then so is its translation $(\phi)^\star$, since our function $(\cdot)^\star$ translates the connectives and quantifiers homophonically.

Consider a proof of length $n+1$, and assume that for $\leqslant n$-long proofs in first order logic of $\psi$ from $\Delta$, $(\Delta)^\star \vdash_{lq} (\psi)^\star$. Then $\Gamma = \Delta \cup \{\psi\}$ and $\phi$ is obtained from $\psi$ by one application of Generalization (Gen), or $\Gamma = \Delta \cup \{\psi, \psi \to \phi\}$ and $\phi$ is obtained from $\psi$ by one application of Modus Ponens (MP).

(MP): $(\Gamma)^\star = (\Delta)^\star \cup \{(\psi)^\star, (\psi \to \phi)^\star\}$ and $(\Gamma)^\star \vdash_{lq} (\phi)^\star$ since *lq* extends classical first order logic and in particular is closed under Modus Ponens.

(Gen): $(\Gamma)^\star = (\Delta)^\star \cup \{(\psi)^\star\}$, and $(\Gamma)^\star \vdash_{lq} (\phi)^\star$ since the translation function leaves quantifiers untouched and *lq* comprises ordinary first order logic. □

The second lemma needed for proving proposition 17 says that translations of KF axioms are MGT theorems.

**Lemma 18.** *Let $\phi$ be a KF axiom. We have that*

$$\mathrm{MGT} \vdash_{\mathrm{lq}} (\phi)^\star$$

*Proof.* By completeness, it suffices to show that if $\phi$ is a KF axiom then

$$\mathrm{MGT} \vDash_{\mathrm{lq}} (\phi)^\star$$

That is, for every linearly ordered, constant domain model $\mathfrak{M} = (W, R, D, d)$ and every point $w \in W$, if $\mathfrak{M} \vDash \mathrm{MGT}[w]$ then $\mathfrak{M} \vDash (\phi)^\star[w]$.

(KF1) We wish to show that $\mathrm{MGT} \vDash_{lq} (\forall x \forall y (Tx \underset{.}{=} y \leftrightarrow x = y))^\star = \forall x \forall y (\diamondtopblack\, Tx \underset{.}{=} y \leftrightarrow x = y)$. ($\leftarrow$) Assume $x = y$. By TKF1, first conjunct, we have that $\diamond, Tx \underset{.}{=} y$. Hence, by definition, $\diamondtopblack\, Tx \underset{.}{=} y$, as desired.

($\rightarrow$) Assume $\diamondtopblack\, Tx \underset{.}{=} y$. There is a point such that $Tx \underset{.}{=} y$. Consequently, $x = y$ at some preceding point. But then, by RT, $\squareulblack\, x = y$ holds at that point, which makes $x = y$ hold at every point, including, by the linearity of $R$, the one we started out from, as desired.

Analogously, we show that $\mathrm{MGT} \vDash_{lq} (\mathrm{KF2})^\star$.

(KF12) We wish to show that $\mathrm{MGT} \vDash_{lq} \forall x (\diamondtopblack\, T\underset{.}{T}x \leftrightarrow \diamondtopblack\, Tx)$. ($\rightarrow$) Assume $\diamondtopblack\, T\underset{.}{T}x$. Hence, at some point, $T\underset{.}{T}x$. By TKF12, $\blacklozenge Tx$ and by definition, $\diamondtopblack\, Tx$. ($\leftarrow$) Assume $\diamondtopblack\, Tx$ and go to the point $\alpha$ where $Tx$. By TKF12, therefore, $T\underset{.}{T}x$ holds at at some later point $\beta$, and we conclude that $\alpha$ witnesses $\diamondtopblack\, T\underset{.}{T}x$.

(KF13) Our goal is to show that $\mathrm{MGT} \vDash_{lq} \forall x (\diamondtopblack\, T\underset{.}{\neg}\underset{.}{T}x \leftrightarrow \diamondtopblack\, T\underset{.}{\neg}(x) \vee \neg Se\underset{.}{n}t_{\mathrm{ta}}(x))$. ($\rightarrow$) As before, we go to a point $\alpha$ that witnesses $\diamondtopblack\, T\underset{.}{\neg}\underset{.}{T}x$, where we use TKF13 to infer $\blacklozenge T\underset{.}{\neg}x$. Moving on to this statement's witness $\beta < \alpha$, we find that here, $T\underset{.}{\neg}x$ holds. We conclude that $\diamondtopblack\, T\underset{.}{\neg}x$ holds at our starting point, as desired. ($\leftarrow$) If $\neg Se\underset{.}{n}t_{\mathrm{ta}}(x))$ or $\diamondtopblack\, T\underset{.}{\neg}x$, we proceed as before, applying TKF13's first conjunct, and conclude $\diamondtopblack\, T\underset{.}{\neg}\underset{.}{T}x$.

(KF3-11) The compositionality axioms are all treated similarly, stripping off $\diamondtopblack$ and at that point, applying the relevant MGT axiom to obtain a witness for the desired claim. For example, to show the right-to-left direction of $\mathrm{MGT} \vDash_{lq} (\mathrm{KF3})^\star$ we go to the witness $\nu$ of $\diamondtopblack\, Tx$. There, we make use of the fact that $T\underset{\cdot}{\neg}\underset{\cdot}{\neg}x \leftrightarrow Tx$ holds at this point, too, to conclude that the point witnesses the desired claim $\Diamond\, T\underset{\cdot}{\neg}\underset{\cdot}{\neg}x$. □

Having lemmata 17 and 18 at hand, I can now prove that MGT interprets KF.

*Proof of proposition 17.* From lemma 18 we know that MGT proves the translations of all KF axioms. Since lemma 17 ensures that derivation in KF is preserved, too, the claim is verified by a simple induction on the length of proof. □

Proposition 17 is a pleasing and useful result. Much is known about the interpretability strength of KF [Halbach, 2011, §15.3]. Proposition 17 allows us to exploit these facts and relate MGT to Tarski's theory of truth and ramified analysis. Let '$RT_{<\alpha}$' denote the theory of Tarskian truth over arithmetic iterated up to the ordinal $\alpha$, and let '$RA_{<\alpha}$' denote the theory of predicative second-order arithmetic, iterated up to the ordinal $\alpha$. Recall that $\epsilon_0$ is the limit of the sequence $\omega, \omega^{\omega}, \omega^{\omega^{\omega}}, \ldots$.

**Corollary 3.** *MGT interprets $RA_{<\epsilon_0}$ and $RT_{<\epsilon_0}$.*

In the precise sense of proposition 17, nothing is lost by couching KF in the logic of linear time. In fact, much is gained. MGT is strictly stronger than KF. KF, on the one hand, does not prove the consistency of truth, more precisely KF$\nvdash \forall x(\mathit{Se\underset{\cdot}{n}t}_{ta}(x) \rightarrow \neg(Tx \wedge T\underset{\cdot}{\neg}x))$. To see this, recall firstly that by Feferman's classic result, for every Strong Kleene fixed point S, $\mathfrak{N}(S)$ is a model of KF, and secondly that there are fixed points that contain both the liar sentence and its negation.

MGT, on the other hand, proves $\Box\, \forall x(\mathit{Se\underset{\cdot}{n}t}_{ta}(x) \rightarrow \neg(Tx \wedge T\underset{\cdot}{\neg}x))$. To see this, we first need to introduce some terminology.

**Definition 28** (T-complexity)**.** Equations $x = y$ have T-complexity 0. The T-complexity of $T\ulcorner\psi\urcorner$ is one greater than the T-complexity of $\psi$. $\neg\phi$ and $\exists x\phi$ inherit the T-rank of $\phi$. The T-complexity of $\phi \wedge \psi$ and $\phi \vee \psi$ is the T-complexity of $\phi$ or $\psi$, whichever greater.

The modal logic of grounded truth proves that truth is always consistent, in the precise sense of the following proposition.

**Proposition 18.** *For every $\mathcal{L}_{ta}$-sentence $\phi$,*

$$\text{MGT} \vdash_{lq} \Box\, \forall x(\text{Se\underset{\cdot}{n}t}_{ta}(x) \rightarrow \neg(Tx \wedge T\underset{\cdot}{\neg}x)$$

*Proof.* For simplicity, I present the proof of the following schema which, however, is emulated within MGT similarly to how lemma 16 was proved.

$$\neg(T\ulcorner\phi\urcorner \wedge T\underset{\cdot}{\neg}\ulcorner\phi\urcorner) \tag{78}$$

Let $\phi$ be any $\mathcal{L}_{ta}$-sentence. By completeness, it suffices to show that $\mathrm{MGT} \models_{lq} \boxslash \neg(T\ulcorner\phi\urcorner \wedge T\ulcorner\neg\phi\urcorner)$. So let $\mathfrak{M}$ be any $lq$ model $(W, R, D, d)$, let $w$ be a point in $W$ and assume that $\mathfrak{M} \models \mathrm{MGT}$. We show that $\mathfrak{M} \models \boxslash \neg(T\ulcorner\phi\urcorner \wedge T\ulcorner\neg\phi\urcorner)[w]$ and do so by an induction on the T-complexity of $\phi$ as just defined. At its base, assuming that $\phi$ is arithmetical, we run an induction on the *syntactic* complexity. So let $\phi$ be an atomic formula of arithmetic, i.e. an equation $a = b$ (we are now at the base of the inner of two nested inductions). Assume, for contradiction, that $T\ulcorner a = b\urcorner \wedge T\ulcorner a \neq b\urcorner$ holds at $w$. Then, by TKF1 and TKF2, $\blacklozenge x = y \wedge \blacklozenge x \neq y$. Hence, at some point $uRv$, $x = y$ and at some point $u'Rv$, $x \neq y$. But by lemma 15, $\boxslash\, x = y \wedge \boxslash\, x \neq y$, contradiction. Hence $\neg(T\ulcorner a = b\urcorner \wedge T\ulcorner a \neq b\urcorner)$ at $v$, as desired.

For complex arithmetical sentences $\phi$, the claim that $\boxslash \neg(T\ulcorner\phi\urcorner \wedge T\underset{\cdot}{\neg}\ulcorner\phi\urcorner)$ follows from the axioms TKF3-11 and the induction hypothesis. For example, $T\ulcorner\neg\psi\urcorner \wedge T\ulcorner\neg\neg\psi\urcorner$ becomes $T\ulcorner\neg\psi\urcorner \wedge T\ulcorner\psi\urcorner$ by TKF3, which directly contradicts our induction hypothesis.

Now assume $\phi$ to be of T-complexity $n+1$, and assume that for all sentences $\psi$ of lower complexity, $\mathfrak{M} \models \neg(T\ulcorner\psi\urcorner \wedge T\ulcorner\neg\psi\urcorner)[w]$. Again, we conduct an induction on the syntactic complexity of $\phi$. If $\phi$ is atomic, we know that $\phi = T\ulcorner\psi\urcorner$ for some sentence $\psi$ of T-complexity $n$. Assume, for contradiction, that $T\ulcorner T\ulcorner\psi\urcorner\urcorner \wedge T\neg T\ulcorner\psi\urcorner$ holds at $w$. By the right-hand conjuncts of TKF12 and TKF13, we know that $\blacklozenge T\ulcorner\psi\urcorner \wedge \blacklozenge T\ulcorner\neg\psi\urcorner$ is true at $w$. Hence, for some $v$ and $u$ both R-earlier than $w$, $T\ulcorner\psi\urcorner$ is true at $v$ and $T\ulcorner\neg\psi\urcorner$ is true at $u$. Since R is a linear ordering, we can assume without loss of generality that $uRv$. By lemma 16 we have that $T\ulcorner\neg\psi\urcorner$ must hold at all points R-later than $u$, in particular at $v$. Hence, $\mathfrak{M} \models T\ulcorner\psi\urcorner \wedge T\ulcorner\neg\psi\urcorner[v]$. Again, since $vRw$ lemma 16 allows us to infer that this conjunction $T\ulcorner\psi\urcorner \wedge T\ulcorner\neg\psi\urcorner$ holds at at the point $w$, contrary to our induction hypothesis.

The induction step, at which we assume $\phi$ to be complex, is taken care of, as before, by the axioms TKF3-11 and the induction hypothesis that for every constituent $\psi$ of $\phi$, $\mathfrak{M} \models \neg(T\ulcorner\psi\urcorner \wedge T\ulcorner\neg\psi\urcorner)[w]$. For example, let $\phi$ be $\exists x\psi$, and assume, for contradiction, that $\mathfrak{M} \models \neg(T\ulcorner\exists x\psi\urcorner \wedge T\ulcorner\neg\exists x\psi\urcorner)[w]$. The axioms TKF10 and TKF11 allow us to infer that $\exists y(ClTm(y) \wedge T\ulcorner\psi(\dot{y}/x)\urcorner) \wedge \forall y(ClTm(y) \rightarrow T\ulcorner\neg\psi(\dot{y}/x)\urcorner)$ must hold at $w$. Let $y_0$ witness the first conjunct, and specialize the second conjunct to it. We get that at $w$, it must be that $\mathfrak{M} \models T\ulcorner\psi(\dot{y}_0/x)\urcorner \wedge T\ulcorner\neg\psi(\dot{y}_0/x)\urcorner[w]$. This, however, contradicts our induction hypothesis.

This completes the proof that for every $lq$-model $\mathfrak{M}$ and every point $w \in W$, if $\mathfrak{M} \models \mathrm{MGT}[w]$ then $(W, R, D, d) \models \boxslash \forall x(\underset{\cdot}{Sent}_{ta}(x) \rightarrow \neg(Tx \wedge T\underset{\cdot}{\neg}x))[w]$. $\Box$

Emulating the corresponding proof for KF+Cons [Halbach, 2011, p. 212] we infer from proposition 18 that according to MGT, only sentences are true.

**Corollary 4.** MGT $\vdash_{lq} \boxbslash\, \forall x(Tx \to \mathrm{S\underset{\cdot}{e}nt}_{ta}(x))$

Theorem 18 is the first piece of evidence that the tense logic setting pays off for the theorist of grounded truth. In addition, it allows her to show that for $\phi$ to be true, it must have been the case that $\phi$ earlier.

**Corollary 5.** *For every $\mathcal{L}_{ta}$-sentence $\phi$,*

$$\mathrm{MGT} \vdash_{lq} \boxbslash\, (T\ulcorner\phi\urcorner \to \blacklozenge\phi)$$

*Proof.* By completeness, it suffices to show for every *lq* model $\mathfrak{M} = (W, R, D, d)$ and every point $w \in W$,

$$\mathfrak{I} \vDash \mathrm{MGT}[w] \Rightarrow \mathfrak{I} \vDash \boxbslash\, (T\ulcorner\phi\urcorner \to \blacklozenge\phi)[w]$$

So let $\phi$ be any $\mathcal{L}_{ta}$-sentence, $w$ some point in $W$ and assume that $\mathfrak{I} \vDash \mathrm{MGT}[w]$. Further, let $v$ be $w$ or any point to the left or right of $w$ (that is, let $v \in W$). In order to show that $\mathfrak{I} \vDash T\ulcorner\phi\urcorner \to \blacklozenge\phi[v]$, we reason by induction on the positive complexity of $\phi$. If $\phi$ atomic then the claim follows directly from TKF1,-2, -12 and TKF13. The interesting case is that of showing $\mathfrak{I} \vDash T\ulcorner\neg Ta\urcorner \to \blacklozenge\neg Ta[v]$. So assume that $\mathfrak{I} \vDash T\ulcorner\neg Ta\urcorner[v]$. Then, since we assume TKF13 to hold at $v$, we know that for some $uRv$, $(W, R, D, d) \vDash T\underset{\cdot}{\neg}a[u]$. At this point, it is theorem 18 and the fact that $\mathfrak{I} \vDash \neg(Ta \wedge T\underset{\cdot}{\neg}a)[u]$ that allows us to proceed and conclude that $\neg Ta$ holds at $u$, hence $\blacklozenge\neg Ta$ holds at $v$.

At the induction step, where we assume $\mathfrak{I} \vDash T\ulcorner\psi\urcorner \to \psi[v]$ to hold for every $\psi$ of lower complexity than $\phi$, the claim follows from a combination of TKF3-11 and the induction hypothesis. For example, assuming $\mathfrak{I} \vDash T\ulcorner\neg(\psi \wedge \zeta)\urcorner[v]$, we infer from $\mathfrak{I} \vDash (\mathrm{TKF5})[v]$ that $T\ulcorner\neg\psi\urcorner$ or $T\ulcorner\neg\zeta\urcorner$ holds at some $uRv$. Either way, however, our induction hypothesis and logic then licences the inference (at $u$) of $\neg(\psi \wedge \phi)$, as desired. $\square$

**Corollary 6.** *For every $\mathcal{L}_{ta}$-sentence $\phi$,*

$$\mathrm{MGT} \vdash_{lq} \boxbslash\, (T\ulcorner\phi\urcorner \to \phi)$$

*Proof.* For every $\phi$, assuming $T\ulcorner\phi\urcorner$ we get $\blacklozenge\phi$ from corollary 5. Then, we show $\phi$ on the basis of lemma 16 or lemma 15, for sentences containing 'T' or not, respectively. $\square$

**Corollary 7.** *Let $\tau$ be a* truth-teller, *such that PA$\vdash \tau \leftrightarrow T\ulcorner\tau\urcorner$. Then*

$$\mathrm{MGT} \vdash_{lq} \boxbslash\, \neg T\ulcorner\tau\urcorner$$

*Proof.* By completeness of *lq*, it suffices to show that for every *lq*-model $\mathfrak{M} = (W, R, D, d)$ and every point $w \in W$, if $\mathfrak{M} \vDash \mathrm{MGT}[w]$ then $\mathfrak{M} \vDash \boxbslash\, \neg T\ulcorner\tau\urcorner[w]$.

So let $\mathfrak{M}$ be an *lq*-model, $w \in W$ and assume that if $\mathfrak{M} \vDash \mathrm{MGT}[w]$. For contradiction, assume that $\mathfrak{M} \vDash \diamondbslash\, T\ulcorner\tau\urcorner[w]$. Then at $w$ or at some point $v$ to the left or right of $w$ (we know R to be linear), $\mathfrak{M} \models T\ulcorner\tau\urcorner[v]$. Since $\mathfrak{M} \models \mathrm{TKF12}[w]$, and $v = w$ or to the left or right of $w$, $\mathfrak{M} \models \diamond T\underset{\cdot}{T}\ulcorner\tau\urcorner[v]$. Hence at some $u$ to the right of $v$, $T\underset{\cdot}{T}\ulcorner\tau\urcorner$. Since $\mathfrak{M} \models \mathrm{TKF12}$ $[w]$ and we know $u$ to be the left or right of, if not identical to $w$, $\mathfrak{M} \models \blacklozenge(T\ulcorner\tau\urcorner \wedge \blacksquare\neg T\ulcorner\tau\urcorner)[u]$. Consequently, at some $t$ left of $u$, $T\ulcorner\tau\urcorner$ as well as $\blacksquare\neg T\ulcorner\tau\urcorner$. By corollary 5, for some further $s$ left of $t$ (hence somewhere to the left or right of, or identical with $w$), $\mathfrak{M} \models \tau[s]$. But because PA holds at $s$ and by our assumption about this sentence $\tau$, $\mathfrak{M} \models T\ulcorner\tau\urcorner[s]$. However, because $\blacksquare\neg T\ulcorner\tau\urcorner$ holds at $t$ to the right of $s$, we also know that $\mathfrak{M} \models \neg T\ulcorner\tau\urcorner[s]$, contradiction. □

Corollary 7 indicates that we are on the right track towards a theory of grounded truth.

It remains an interesting question how MGT relates to Burgess' theory KFB. As explained in section 4.3 (p. 70), KFB is intended as an axiomatization of the least fixed point, and likewise proves truth to be consistent and a truth-teller not to be true. One thing is clear, though. We want MGT to do better than KFB. As recently observed by Volker Halbach, KFB holds in other fixed points than the least one [Fischer et al., 2014]. Thus, KFB is not capable of singling out exactly the grounded truths.

The results of this section provide some evidence that MGT, unlike KF, is a theory of grounded truth. The main challenge, however, is to show that our theory can express enough of the original, semantic notion of groundedness. In the next section, I will make first steps into this direction.

## 9.6 MGT AND THE STAGES OF KRIPKE'S CONSTRUCTION

I now turn MGT's semantics. My goal in this section is to argue that MGT is a theory of *grounded* truth in a very robust sense. The main result of this section is proposition 19. It shows that MGT has a natural model: the stages of Kripke's construction (see §3.3).[6] Moreover, MGT characterizes this particular model exactly.

How can this be? After all, the Kripke stages are well-ordered. As we saw in section 9.4 (corollary 1), the logic of well-ordered time is not axiomatizable. In particular, MGT is based on a logic of linearly ordered time. Therefore, the theory MGT cannot distinguish between a Kripke-like but ill-founded sequence of models of truth, and my goal, the real order of Kripke stages.

As much as this is true, however, it is also orthogonal to the question whether MGT characterizes the Kripke construction. Let me give an analogy. KF is generally considered an adequate axiomatization of the Strong Kleene fixed points. However, it is based on a merely

6 For simplicity, I will in this section work with the standard, *low-resolution* presentation of Kripke's construction (§3.3).

first-order theory of arithmetic, whereas a Kripke fixed point is an expansion of the *standard* model. And of course, first-order Peano Arithmetic cannot single out the standard model. Nonetheless KF is considered adequate. The reason is that *assuming arithmetic to be standard*, we can show that KF singles out the fixed points [Halbach, 2011, theorem 15.15]. Formally, for all sets X of $\mathcal{L}_{ta}$-sentences,

$$\mathfrak{N}(X) \models \mathrm{KF} \Leftrightarrow X = \mathcal{J}_{sk}(X) \tag{79}$$

My goal is to show that MGT characterizes the stages of Kripke's construction just as well. *Assuming time to be well-ordered*, we can show that MGT characterizes the stages. More precisely, I will show that *well-ordered* models of MGT are isomorphic to the stages of Kripke's construction.

It may be thought that now I make too many assumption for the result to have much significance. Since, clearly, I still have to assume the number to be standard, as in the non-modal case. Thus, the theory characterizes groundedness only within the doubly narrow range of well-ordered, standard models.

However, things are not as they seem. Recall proposition 1 (p. 163). There is a set of principles N1-N3 in the language of tensed first-order arithmetic, such that any well-ordered model $\mathfrak{M}$ validates first-order arithmetic plus N1-N3 only if $\mathfrak{M}$ interprets the arithmetical vocabulary in the standard model $\mathfrak{N}$. I will show that MGT proves, in linear time, such a set of principles (lemma 21). Hence, every well-ordered model of MGT respects the theory of standard numbers. Consequently, the analogy between the adequacy of KF of my result 19 is robust. In fact, just as for KF we *only* assume standardness, I now *only* have to assume time to be well-ordered. Making this assumption, we will get the standard numbers for free.

To show that MGT proves principles that characterize the natural numbers, some preparation is needed. Firstly, observe that the syntactic relation "the formula $\phi$ is the result of applying $x$ iterations of 'T' to $\psi$" is represented in PA by an $\mathcal{L}_{ta}$-formula that I will denote $T^{x}\ulcorner\phi\urcorner$, such that PA$\vdash$ $\ulcorner\phi\urcorner = T^{x}\ulcorner\psi\urcorner$ iff, roughly,

$$\phi = \text{“}\underbrace{\mathrm{T}\dot{\mathrm{T}}\ldots\dot{\mathrm{T}}}_{x\text{-many}}\ulcorner\psi\urcorner\text{”}$$

Using the $\dot{x}$ function from chapter 4 (p. 47) we thus enable the theory to quantify into the argument place $x$.

**Lemma 19.**

$$\mathrm{MGT} \models_{lq} \;\blacksquare\; \forall u \forall y \forall z \forall x \big(x = y + z \rightarrow (\mathrm{T}T^{\dot{x}}u \rightarrow \mathrm{T}T^{\dot{y}}u)\big)$$

*Proof.* We reason within MGT by (first-order) induction on $z$. The base case where $x = y$ is a truth of logic. The induction step follows from the first conjunct of axiom TKF12 and lemma 16, using the induction hypothesis. □

Secondly, let us define a special infinite sequence of $\mathcal{L}_{\mathrm{ta}}$-formulae.

**Definition 29.** $\theta(x) := \mathrm{T}\mathrm{T}^{x}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg\mathrm{T}\mathrm{T}^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner$

For example, $\theta(\overline{0})$ is the sentence $\mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg\mathrm{T}\dot{\mathrm{T}}\ulcorner\overline{0} = \overline{0}\urcorner$.[7] The open formula $\theta(x)$ of the language of tensed truth $\mathcal{L}_{\mathrm{tam}}$ will play the role of the predicate 'N' in terms of which we formulated the principles N1-N3 of proposition 1. In order to show this, however, one further lemma is needed.

**Lemma 20.** *For every* lq*-model* $\mathfrak{M} = (W, R, D, d)$ *and every point* $w \in W$*, if* $\mathfrak{M} \models$*MGT*$[w]$ *then there is a point at which nothing is true but at every point accessible from it, something is true; more precisely, there is a* $v \in W$ *such that* $d(v)(\text{'T'}) = \varnothing$ *and for every* $u \in W$ *such that* $vRu$*,* $d(u)(\text{'T'}) \neq \varnothing$*.*

*Furthermore, this point* $v$ *is the* least *point in the linear order* $R$*: there is no point* $u$ *which sees* $v$*.*

*Proof.* Let $\mathfrak{M}$ be any linearly ordered constant-domain model, and $w$ any point in it. Assume that $\mathfrak{M} \models$MGT$[w]$. Then $\mathfrak{M}$ validates the axiom of Ground: $\mathfrak{M} \models \text{⬘}\forall x\neg Tx[w]$, i.e. $\mathfrak{M} \models \blacklozenge\forall x\neg Tx \vee \forall x\neg Tx \vee \lozenge\forall x\neg Tx[w]$. In each of these cases, there is some $v$ such that $\mathfrak{M} \models \forall x\neg Tx[v]$.

Now, by first order logic and because $\mathfrak{M}$ validates MGT, in particular TKF1, $\mathfrak{M} \models \text{◩}(\overline{0} = \overline{0}) \wedge \text{◩}(\overline{0} = \overline{0} \rightarrow \Box\mathrm{T}\ulcorner\overline{0} = \overline{0})\urcorner[w]$.

By the linearity of $R$ we know that $wRv$, $vRw$ or $v = w$. In any case, $\mathfrak{M} \models \overline{0} = \overline{0} \wedge \overline{0} = \overline{0} \rightarrow \Box\mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner[v]$. Hence, $\mathfrak{M} \models \Box\mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner[v]$ and at every point $u$ accessible from $v$, $\mathfrak{M} \models \mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner[u]$, hence $d(u)(\text{'T'}) \neq \varnothing$.

It remains to show that there is no point $u$ which sees $v$. For contradiction, assume that there is such a point $u$. Then, for the same reason as before, $\mathfrak{M} \models \Box\mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner[u]$. But since $uRv$, this means that $\mathfrak{M} \models \mathrm{T}\ulcorner\overline{0} = \overline{0}\urcorner[v]$, contradiction. $\square$

Finally, we are now in a position to show that MGT proves principles of the kind which we know to require a standard interpretation of the natural numbers (principles N1 to N3 on p. 163).

**Lemma 21.** *Let* $\theta(x)$ *be the formula as defined in* 29*. We have that in the logic of linear time, MGT proves the following principles.*

θ.1 $\forall x \, \text{⬘} \big(\theta(x) \wedge \blacksquare\neg\theta(x) \wedge \Box\neg\theta(x)\big)$

θ.2 $\text{◩} \, \forall x\forall y\big((\theta(x) \wedge \theta(y)) \rightarrow x = y\big)$

θ.3 $\text{◩} \, \forall x\forall y\Big(x = y + 1 \leftrightarrow \text{⬘}\big(\theta(y) \wedge \lozenge\theta(x)\big) \wedge \forall z\big(x \neq z \wedge \text{⬘}(\theta(y) \wedge \lozenge\theta(z)) \rightarrow \text{⬘}(\theta(x) \wedge \lozenge\theta(z))\big)\Big)$

7 Note that we cannot define $\theta(x)$ as $\mathrm{T}^{x}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg\mathrm{T}^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner$ since $\mathrm{T}^{x}\ulcorner\overline{0} = \overline{0}\urcorner$ is not a sentence, but a term.

*Proof.* (θ.1) In order not to assume standard numbers from the outside we reason within MGT, by induction on x. For this, we need to show that

$$\begin{aligned}\mathrm{MGT} \vdash_{lq} &\ \diamond\!\!\!\!\cdot\ \big(\theta(\overline{0}) \wedge \blacksquare\neg\theta(\overline{0}) \wedge \square\neg\theta(\overline{0})\big)\\ &\wedge \forall x\Big(\forall y\big(x < y \rightarrow\\ &\big(\ \diamond\ \big(\theta(\overline{y}) \wedge \blacksquare\neg\theta(\overline{y}) \wedge \square\neg\theta(\overline{y})\big) \rightarrow \diamond\ \big(\theta(\overline{x}) \wedge \blacksquare\neg\theta(\overline{x}) \wedge \square\neg\theta(\overline{x})\big)\big)\big)\Big)\end{aligned}$$

For the first conjunct, by the completeness of *lq* it suffices to show that for every *lq* model $\mathfrak{M}$, and every point $w$, if $\mathfrak{M} \models$MGT$[w]$ then there is some $v$ to the left or right of $w$ such that

$$\mathfrak{M} \models T\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg T\underset{\cdot}{T}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \blacksquare\neg\theta(\overline{0}) \wedge \square\neg\theta(\overline{0})[v] \tag{80}$$

So let $w$ be any point in $W$ and assume $\mathfrak{M} \models$MGT$[w]$. From lemma 20, we know that there is an R-least point $w_0$ such that $\mathfrak{M} \models \neg T\ulcorner\overline{0} = \overline{0}\urcorner \wedge \overline{0} = \overline{0}[w_0]$. By the second conjunct and axiom TKF1, there is a point $v$, $w_0 R v$, such that $\mathfrak{M} \models T\ulcorner\overline{0} = \overline{0}\urcorner[v]$. By the second conjunct of axiom TKF12, then, we know that there is a point $u$, $vRu$, such that $\mathfrak{M} \models T\underset{\cdot}{T}\ulcorner\overline{0} = \overline{0}\urcorner[u]$. Now we make use of the first conjunct of TKF12, and infer that there must be a point $u'$, $u'Ru$, at which $T\ulcorner\overline{0} = \overline{0}\urcorner \wedge \blacksquare\neg T\ulcorner\overline{0} = \overline{0}\urcorner$ is true. Let this be our witness. Firstly, by the first conjunct and axiom TKF12 we have that $\mathfrak{M} \models \square T\underset{\cdot}{T}\ulcorner\overline{0} = \overline{0}\urcorner[u']$, hence $\square\neg\theta(\overline{0})$. Secondly, by the second conjunct we know that at no point to the left of $u'$ $T\ulcorner\overline{0} = \overline{0}\urcorner$ will be true, hence $\mathfrak{M} \models \blacksquare\neg\theta(\overline{0})[u']$. Finally assume, for contradiction, that $\mathfrak{M} \not\models \neg T\underset{\cdot}{T}\ulcorner\overline{0} = \overline{0}\urcorner[u']$. Then $T\underset{\cdot}{T}\ulcorner\overline{0} = \overline{0}\urcorner$ must be true at this point, hence $\blacklozenge T\ulcorner\overline{0} = \overline{0}\urcorner$. But we already know that $\blacksquare\neg T\ulcorner\overline{0} = \overline{0}\urcorner$ is true at $u'$, contradiction. Therefore $\mathfrak{M} \models \theta(\overline{0})[u']$, as desired.

At the induction step, it again suffices to show that for every *lq* model $\mathfrak{M}$ and point $w$, for every $o \in D$, if

$$\mathfrak{M} \models \mathrm{MGT} \wedge \forall y\big(y < x \rightarrow \big(\ \diamond\ \big(\theta(y) \wedge \blacksquare\neg\theta(y) \wedge \square\neg\theta(y)\big)\big)[x : o][w]$$

then

$$\mathfrak{M} \models\ \diamond\ \big(\theta(x) \wedge \blacksquare\neg\theta(x) \wedge \square\neg\theta(x)\big)[x : o][w]$$

So let $p$ be any object from the domain D, and assume $\mathfrak{M} \models x = y + 1[x : o][y : p][w]$, such that for some point $v$ to the left of right of $w$, $TT^y\ulcorner\overline{0} = \overline{0}\urcorner$ is true at $v$ for $o$ assigned to $x$ and $p$ assigned to $y$.[8] Twice making use of the second conjunct of axiom TKF12, we have that for some $u$ seen by $v$, $\mathfrak{M} \models T\underset{\cdot}{T}T^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][u]$. Then, by the axiom's first conjunct we know that at some $u'$, $u'Ru$, it is true that $TT^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \blacksquare\neg TT^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner$. Recall that $\mathfrak{M} \models x = y + 1[x : o][y : p]$. Analogously to before, we therefore let this $u'$ be our witness.

8 I will suppress the variable assignment where possible.

On the one hand, by the second conjunct we know that at $u'$, $\blacksquare\neg TT^x\ulcorner\overline{0} = \overline{0}\urcorner$, hence $\blacksquare\neg\theta(x)$. From the first conjunct and axiom TKF12 we know that $\mathfrak{M} \models \Box T\underset{\cdot}{T}T^x\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][u']$, hence $\Box\neg\theta(x)$ is true at $u'$, too. Recall that $\theta(x)$ is the formula $TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner$.

On the other hand, assume for contradiction that $\mathfrak{M} \models TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][u']$. Then at $u'$, it must be true that $\blacklozenge TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \blacksquare\neg TT^x\ulcorner\overline{0} = \overline{0}\urcorner$, contradiction.

($\theta$.2) follows from lemma 19 since, if by contraposition and without loss of generality we assume that $x < y$ then $\theta(y)$ entails $TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner$, hence $\neg\theta(x)$.

($\theta$.3.) By completeness it suffices to show that for every *lq* model $\mathfrak{M} = (W, R, D, d)$ and every point $w \in W$, if $\mathfrak{M} \models$MGT$[w]$ then for every $v \in W$ and every $o, p \in D$, $\mathfrak{M} \models x = y + 1[x : o][y : p][v]$ iff

i. for some $u$, $\mathfrak{M} \models \theta(y) \wedge \Diamond\theta(x)[x : o][y : p][u]$

ii. $\mathfrak{M} \models \forall z\big(x \neq z \wedge \mathbin{\diamondsuit}(\theta(y) \wedge \Diamond\theta(z)) \rightarrow \mathbin{\diamondsuit}(\theta(x) \wedge \Diamond\theta(z))\big)[x : o][y : p][v]$

For the left-to-right direction, assume $\mathfrak{M} \models x = y + 1[x : o][y : p][v]$. For (i) note that by ($\theta$.1), there are $u$ and $u'$ such that

$$\mathfrak{M} \models TT^y\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^x\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][u] \tag{81}$$

$$\mathfrak{M} \models TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][u'] \tag{82}$$

To show that $uRu'$, by the linearity of $R$ it suffices to note that $u$ and $u'$ cannot be identical, and that since by lemma 16, $\mathfrak{M} \models \Box TT^x[x : o][y : p][u']$ assuming $u'Ru$ leads equally to contradiction.

For (ii), let $r \in D$ such that $\mathfrak{M} \models z \neq x[x : o][y : p][z : r][v]$ and assume that for some $u, u'$, $uRu'$, (81) and

$$\mathfrak{M} \models TT^z\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{z+1}\ulcorner\overline{0} = \overline{0}\urcorner[x : o][y : p][z : r][u'] \tag{83}$$

By ($\theta$.1) we know that there is a $u''$ such that

$$\begin{aligned}\mathfrak{M} \models\, & TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner \\ & \wedge \blacksquare\neg(TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner) \\ & \wedge \Box\neg(TT^x\ulcorner\overline{0} = \overline{0}\urcorner \wedge \neg TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner)[x : o][y : p][z : r][u'']\end{aligned} \tag{84}$$

By lemma 19 we can infer from this that $uRu''$. It remains to show that $u''Ru'$. On the one hand, we note that if $u' = u''$ and $\mathfrak{M} \models z > x[x : o][y : p][z : r][u'']$ then, since $TT^z\ulcorner\overline{0} = \overline{0}\urcorner$ is true at $u''$, lemma 19 implies that $TT^{x+1}\ulcorner\overline{0} = \overline{0}\urcorner$ must also be true there, thus contradicting (84). For $\mathfrak{M} \models z < x[x : o][y : p][z : r][u'']$ the dual argument leads to a contradiction. On the other hand, it likewise cannot be the case that $u'Ru''$ since if $\mathfrak{M} \models x < z[x : o][y : p][z : r][u'']$ then lemma 19 requires $TT^x\ulcorner\overline{0} = \overline{0}\urcorner$ to be true at $u'$. By (84), however, $\mathfrak{M} \models \neg TT^x\ulcorner\overline{0} = \overline{0}\urcorner[u']$, contradiction. Again, by the dual argument we also rule out the case in which $\mathfrak{M} \models z < x[x : o][y : p][z : r][u'']$.

For the right-to-left direction, we firstly note that by lemma 16, any $o, p$ satisfy $\theta(x) \wedge \diamond\theta(y)$ at $v$ only if they satisfy $\neg TT^y$ there, too. Now, assume (i) and (ii) and, for contradiction, that $o$ and $p$ satisfy $x \neq y+1$ at $v$. From (i) we get that $\mathfrak{M} \blacklozenge (\theta(x) \wedge \diamond\theta(y))[x:o][y:p][v]$. I claim that $o$ and $p$ satisfy $\blacklozenge(\theta(x) \wedge \diamond\theta(y+1))$ at $v$.

Now, assume that $o$ and $p$ satisfy $\blacklozenge(\theta(x) \wedge \diamond\theta(y+1))$ at $v$ and $\mathfrak{M} \vDash x > y+1[x:o][y:p][v]$. By our first observation above we then have that at $v$, $o$ and $p$ satisfy $\neg TT^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner$. However, since they satisfy $TT^x\ulcorner\overline{0} = \overline{0}\urcorner$ and $x > y+1$ at $v$, lemma 19 requires that $\mathfrak{M} \vDash TT^{y+1}[x:o][y:p][v]$ after all, contradiction.

If $o$ and $p$ satisfy $x < y+1$ at $v$ then by our observation above and the assumption that $\mathfrak{M} \vDash \blacklozenge(\theta(x) \wedge \diamond\theta(y+1))[x:o][y:p][v]$ $\mathfrak{M} \vDash \neg TT^x\ulcorner\overline{0} = \overline{0}\urcorner[x:o][y:p][v]$, contradicting, once more, lemma 19 according to which at $u$, $o$ and $p$ satisfy $TT^y\ulcorner\overline{0} = \overline{0}\urcorner \wedge (TT^y\ulcorner\overline{0} = \overline{0}\urcorner \rightarrow TT^x\ulcorner\overline{0} = \overline{0}\urcorner)$.

It remains to show my claim that $o$ and $p$ satisfy $\blacklozenge(\theta(x) \wedge \diamond\theta(y+1))$ at $v$, i.e. that for some point $u$, $\mathfrak{M} \vDash \theta(y) \wedge \diamond\theta(y+1)[x:o][y:p][u]$. By $\theta.1$ we know that there is a $u$ such that $\mathfrak{M} \vDash TT^y\ulcorner\overline{0} = \overline{0}\urcorner[x:o][y:p][u]$. Making twice use of the second conjunct of axiom TKF12, we know that there is a $u'$, $uRu'$ such that $o, p$ satisfy $T\dot{T}T^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner$ at $u'$. By the axiom's first conjunct we then know that for some $u''$ R-between $u$ and $u'$, $\mathfrak{M} \vDash TT^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner \wedge \blacksquare\neg TT^{y+1}\ulcorner\overline{0} = \overline{0}\urcorner[x:o][y:p][u'']$. By the reasoning as used in the proof of $\theta.1$ we show that in fact, $p$ satisfies $\theta(y+1)$ at this $u''$, which thus witnesses the truth of $\diamond\theta(y+1)$ at $u$, such that we can conclude that $o, p$ satisfy $\blacklozenge(\theta(x) \wedge \diamond\theta(y+1))$ at $u$. □

**Lemma 22.** *For every* woq-*model* $\mathfrak{M} = (W, R, D, d)$, *if*

$$\forall w \in W\,(W, R, D, d) \vDash \mathrm{MGT}[w]$$

*then* $\mathfrak{M}$ *interprets the arithmetical vocabulary standardly.*

*Proof.* From theorem 1 and the previous lemma, which shows that MGT provides us with precisely such a set of principles that characterizes, over a well-ordered frame and together with Robinson arithmetic, the natural numbers. □

Recall Kripke's construction of an $\mathcal{L}_{ta}$ model, based on Strong Kleene logic.

For the purpose of axiomatizing Kripke's theory of truth, it is common to work with the *closed off* fixed point model $\mathfrak{N}(I^+_{sk})$. However, we may also consider closing off each stage of Kripke's construction. Thus, we arrive at a well-ordering of (classical) models $\mathfrak{N}(I^{+,\alpha}_{sk})$, which gives naturally rise to a model for the modal logic *lq*.[9]

9 Recall that Kripke's construction closes off at the least non-constructive ordinal $\omega_1^{CK}$.

**Definition 30.** Recall that the least non-constructive ordinal $\omega_1^{CK}$ is the set of all constructive ordinals. Let $d_k$ map each constructive ordinal $< \omega_1^{CK}$ to a model of the language $\mathcal{L}_{tam}$ such that firstly, at every point, the arithmetical vocabulary is interpreted in the standard way on the set of natural numbers $\omega$, and secondly, $d_k(\alpha)(\text{'T'}) = \mathfrak{N}(I_{sk}^{+,\alpha})$.
Let KS be the model $(\omega_1^{CK}, <, \omega, d_k)$.

Recall that I write $S \simeq S'$ if the structure S is isomorphic to $S'$. The main result of this chapter is that all natural models of MGT, in the precise sense of the following proposition, are isomorphic to Kripke's stages.

**Proposition 19.** *For every* woq-*model* $\mathfrak{M} \models (W, R, D, d)$,

$$\forall w \in W\, \mathfrak{M} \models \text{MGT}[w] \textit{ if and only if } (W, R, D, d) \simeq \text{KS}$$

*Proof.* ($\Rightarrow$) By lemma 22 we know that $\mathfrak{M}$ interprets the arithmetical vocabulary standardly, such that we can, for simplicity, identify every point of $W$ with a model $\mathfrak{N}(X)$.

Having noted this, we reason by induction on the well-ordering R. We show that the R-least model is $\mathfrak{N}(\varnothing) = \mathfrak{N}(I_{sk}^{+,0})$. Then, assuming that some point $w$ is the stage $\mathfrak{N}(I_{sk}^{+,\alpha})$, we show that the R-next point $v$ is the successor stage $\mathfrak{N}(I_{sk}^{+,\alpha+1})$. Finally, we show that the R-limit of an initial segment of the points, which we know are the models $\mathfrak{N}\left(I_{sk}^{+,\gamma}\right)$ for $\gamma < \beta$, is the union model $\mathfrak{N}(\bigcup\limits_{\gamma<\beta} I_{sk}^{+,\gamma})$.

So, let $w_0$ be the R-least model $\mathfrak{N}(X)$ in $W$. Since $\mathfrak{M} \models$(Ground)$[w_0]$, $w_0$ or some point R-earlier than $w_0$ must be a model $\mathfrak{N}(\varnothing)$. But there is no such point – after all, $w_0$ is the R-least point. Therefore, it must be that $w_0 = \mathfrak{N}(\varnothing)$.

Now, assume that $w = \mathfrak{N}(I_{sk}^{+,\alpha})$. We need to show that $w$'s R-successor $v$ is $\mathfrak{N}(I_{sk}^{+,\alpha+1})$. We know that $v = \mathfrak{N}(X)$. It remains to show, therefore, that $X = \{\ulcorner\phi\urcorner : \mathfrak{N}(I_{sk}^{+,\alpha}, I_{sk}^{-,\alpha}) \models_{sk} \phi\}$.

($\subseteq$) Let $\ulcorner\phi\urcorner \in X$, we want to show that $\mathfrak{N}(I_{sk}^{+,\alpha}, I_{sk}^{-,\alpha}) \models_{sk} \phi$. We know that $v \models T\ulcorner\phi\urcorner$ and reason by induction on the *positive complexity* of $\phi$. If $\phi$ ='$a = b$' then $v \models \blacklozenge a = b$, since we assume axiom TKF1 to hold at $v$. Hence, at some point $uRv$, $a = b$. Note that since $a = b$ does not involve a partial predicate, if $a = b$ holds in some classical model $\mathfrak{N}(X^+)$ then it also holds in the partial model $\mathfrak{N}(X^+, X^-)$. Therefore, if $u = \mathfrak{N}(I_{sk}^{+,\alpha})$ we are already done. If not, we make use of lemma 15 and conclude that $a = b$ must hold at every point; in particular, therefore, $\mathfrak{N}(I_{sk}^{+,\alpha}) \models a = b$. Hence $\mathfrak{N}(I_{sk}^{+,\alpha}, I_{sk}^{-,\alpha}) \models_{sk} a = b$, as desired.

For $\phi$ ='$a \neq b$' we reason just analogously, using TKF2 instead of TKF1.

If $\phi$ ='Tb' such that $\mathfrak{M} \models T\ulcorner Tb\urcorner[v]$ then we know by the second conjunct of TKF12 that $\mathfrak{M} \models \blacklozenge Tb[v]$. Hence, for some $uRv$, $u \models Tb$.

If $u = w = \mathfrak{N}(I^{+,\alpha}_{sk})$ then we are done. If $uRw$, then lemma 2 allows us to infer that $w \models Tb$ after all, and in fact $\mathfrak{N}(I^{+,\alpha}_{sk}, I^{-,\alpha}_{sk}) \models_{sk} Tb$, as desired.

If $\phi =$'$\neg Tb$' then we use the second conjunct of TKF13. This time, we get that $\mathfrak{M} \models \blacklozenge T\dot{\neg} b \vee \neg \mathit{Sent}_{ta}(s)[v]$. Assume that b denotes a sentence code $\ulcorner\psi\urcorner$ in $v$. Then we know, as before, that $\mathfrak{N}(I^{+,\alpha}_{sk}) \models_{sk} T\ulcorner\neg\psi\urcorner$. If b does not denote a sentence code, then we know that its denotation is in the anti-extension $I^{-,\alpha}_{sk}$, indeed has been so from the first point onwards, and we conclude that $\mathfrak{N}(I^{+,\alpha}_{sk}, I^{-,\alpha}_{sk}) \models_{sk} \neg Tb$, as desired.

Now, consider $\phi =$'$\psi \wedge \zeta$' such that $v = \mathfrak{N}(X) \models T\ulcorner\psi \wedge \zeta\urcorner$. Since $\mathfrak{M} \models$TKF4$[v]$, we know that $v \models T\ulcorner\psi\urcorner \wedge T\ulcorner\zeta\urcorner$. Hence, $\ulcorner\psi\urcorner, \ulcorner\zeta\urcorner \in X$. By our induction hypothesis, we know that $\mathfrak{N}(I^{+,\alpha}_{sk}, I^{-,\alpha}_{sk}) \models_{sk} \psi \wedge \zeta$, as desired. For $\phi =$'$\neg(\psi \wedge \zeta)$' we proceed analogously, exploiting the fact that TKF5 holds in the model.

Disjunction and the quantifiers are taken care of analogously. Recall that $\rightarrow$ is defined in terms of $\neg$ and $\vee$.

($\supseteq$) Let $\mathfrak{N}(I^{+,\alpha}_{sk}, I^{-,\alpha}_{sk}) \models_{sk} \phi$, we want to show that $\ulcorner\phi\urcorner \in X$, that is, $\mathfrak{M} \models T\ulcorner\phi\urcorner[v]$. Since partial truth in a model is contained by classical truth in it, we have that $\mathfrak{N}(I^{+,\alpha}_{sk}) \models \phi$. We reason by induction on the positive complexity of $\phi$. If $\phi =$'$a = b$' then $\mathfrak{N}(I^{+,\alpha}_{sk}) \models \Box T\ulcorner a = b\urcorner$, since we assume TKF1 to hold at $v = \mathfrak{N}(I^{+,\alpha}_{sk})$. We assume $v$ to be the R-successor of $w = \mathfrak{N}(I^{+,\alpha}_{sk})$, hence $\mathfrak{M} \models T\ulcorner a = b\urcorner[w]$, as desired. Analogously for negated equations and sentences of the form Tb or $\neg$Tb.

If $\phi =$'$\psi \wedge \zeta$' we know from our induction hypothesis that $T\ulcorner\psi\urcorner \wedge T\ulcorner\zeta\urcorner$ is true in $\mathfrak{N}(I^{+,\alpha+1}_{sk})$. From the fact that $\mathfrak{N}(I^{+,\alpha+1}_{sk}) \models$TKF4 we further know that $\mathfrak{N}(I^{+,\alpha+1}_{sk}) \models T\ulcorner\psi \wedge \zeta\urcorner$, as desired. Analogously for negated conjunctions and the other connectives and quantifiers. Consequently, $X = \{\ulcorner\phi\urcorner : \mathfrak{N}(I^{+,\alpha}_{sk}, I^{-,\alpha}_{sk}) \models_{sk} \phi\}$, and $v = \mathfrak{N}(I^{+,\alpha+1}_{sk})$, as desired.

Finally, assume that $w_0, \ldots, v$ are the points $\mathfrak{N}\left(I^{+,\alpha}_{sk}\right)$ for $\alpha < \beta$. We show that the R-limit of the points is the model $\mathfrak{N}(\bigcup_{\gamma<\beta} I^{+,\gamma}_{sk})$. We know that $v = \mathfrak{N}(X)$ and show that $X = \bigcup_{\gamma<\beta} I^{+,\gamma}_{sk}$. The reasoning is similar to before and I confine myself to an outline. We reason by induction on the positive complexity of $\phi$. For ($\subseteq$), we observe that for atomic sentences, the axioms TKF1, -2, -12 and -13 ensure that if $\ulcorner\phi\urcorner \in X$ then it is true at some preceding point, hence in their union $\bigcup_{\gamma<\beta} I^{+,\gamma}_{sk}$. The same can be inferred for complex sentences, from the induction hypothesis and the compositionality axioms TKF3-11.

($\Leftarrow$) Let $w$ be any $\mathfrak{N}(I^{+,\alpha}_{sk})$. Of course, KS$\models$ PA$[w]$. The truth of the axiom (Ground) is witnessed by the point $\mathfrak{N}(\varnothing)$. To see that the modal axioms TKF1, -2 and -12 are sound, firstly note that generally, $\mathfrak{N}(I^{+,\alpha}_{sk}) \models T\ulcorner\phi\urcorner$ only if for some $\beta < \alpha$, $\mathfrak{N}(I^{+,\beta}_{sk}) \models \phi$. This validates the right-hand conjuncts.

I now turn to the left-hand conjuncts of axioms TKF1, -2 and -12. Let $\alpha$ be any stage. If atomic sentences $a = b$ and $Ta$ hold in the classical model $\mathfrak{N}(I_{sk}^{+,\alpha})$, then so they do in the partial model $\mathfrak{N}(I_{sk}^{+,\alpha}, I_{sk}^{-,\alpha})$. Hence $\mathfrak{N}(I_{sk}^{+,\alpha+1}) \models T\ulcorner a = b\urcorner \wedge T\ulcorner Tc\urcorner$. Further, since the sequence $(I_{sk}^{+,\alpha})_\alpha$ increases, $T\ulcorner a = b\urcorner \wedge T\ulcorner Tc\urcorner$ will hold at every later stage $\mathfrak{N}(I_{sk}^{+,\beta})$, $\beta > \alpha$. This, however, is what the first conjuncts of TKF1 and TKF12 say, which are thereby shown to hold at every point in the model KS. TKF2 is taken care of analogously, noting that for arithmetical sentence of the form $a \neq b$, classical and partial truth at a stage coincide.

TKF13 requires a subtler treatment, since it concerns *negated* truth, whose behaviour on classical models differs from that on partial ones. Note, however, that the complication is not due to my modal setting, but pertains to the fact that we axiomatize Kripkean truth classically, and hence applies already to KF itself. As a result, the considerations which show the soundness of KF in any Strong Kleene fixed point can guide our investigation into the soundness of MGT. Generally, in order to see the soundness of TKF13, recall that $\neg\phi$ is in some extension $I_{sk}^{+,\alpha}$ just in case $\phi$ is in the corresponding *anti*-extension $I_{sk}^{-,\alpha}$, such that $T\ulcorner\neg\phi\urcorner$ again is in the successor extension $I_{sk}^{+,\alpha+1}$; and that for every term $a$ not standing for a sentence code, $\neg Ta$ is in the extension of 'T' right from the beginning.

In order to show that the compositionality axioms TKF3-11 hold in the model $\mathfrak{N}(I_{sk}^{+,\alpha})$ it suffices to note that the truths in a Strong Kleene model are closed under double-negation, conjunction, disjunction and quantification; and further, that a negated conjunction (disjunction) is classically true just in case both of (one of) the negated conjuncts (disjuncts) are.[10]

This completes the proof of proposition 19. □

Recall that $I_{sk}^{+}$ is the extension of the Strong Kleene fixed point – the set of *grounded* truths. Recall further that we write $\Sigma \models_{woq} \phi$ if $\phi$ is a consequence of $\Sigma$ over the well-ordered frames.

**Corollary 8.** *For every $\mathcal{L}_{ta}$-sentence $\phi$,*

$$\ulcorner\phi\urcorner \in I_{sk}^{+} \textit{ iff } \mathrm{MGT} \models_{woq} \Diamond T\ulcorner\phi\urcorner$$

*Proof.* We show that $\ulcorner\phi\urcorner \in I_{sk}^{+}$ just in case: for every *woq*-model $\mathfrak{M} = (W, R, D, d)$, and for every point $w \in W$ if $\mathfrak{M} \models \mathrm{MGT}[w]$ then $\mathfrak{M} \models \Diamond T\ulcorner\phi\urcorner[w]$.

($\Leftarrow$) Assume that for every *woq*-model $\mathfrak{M} = (W, R, D, d)$ and every point $w \in W$, $\mathfrak{M} \models \Diamond T\ulcorner\phi\urcorner[w]$ if $\mathfrak{M} \models \mathrm{MGT}[w]$. We know, from the right-to-left direction of proposition 19, that the stages of Kripke's construction model MGT: for every $\alpha < \omega_1^{CK}$, KS$\models$ MGT$[\alpha]$. By our

10 To see this, it is instructive to consult Halbach's lemma 15.6 and surrounding remarks [2011, p. 205].

assumption, therefore, for some point $\beta$, KS$\models \blacklozenge$ $T\ulcorner\phi\urcorner[\beta]$. Hence, $\phi \in d_k(\beta) = \mathfrak{N}(I^{+,\beta}_{sk})$, for some $\beta$. Consequently, $\ulcorner\phi\urcorner \in I^{+}_{sk}$, as desired.

($\Rightarrow$) Now assume $\ulcorner\phi\urcorner \in I^{+}_{sk}$, such that $\ulcorner\phi\urcorner \in I^{+,\alpha}_{sk}$, for some $\alpha$. From the left-to-right direction of proposition 19 we know that for every *woq*-model $(W, R, D, d)$ MGT is true at every point in $W$ only if the model is KS. Hence, trivially, in every model that models MGT there is a point, namely $\mathfrak{N}(I^{+,\alpha}_{sk})$, such that $T\ulcorner\phi\urcorner$ holds at this point. Hence, for every *woq*-model and every point, if MGT holds there, so does $\blacklozenge$ $T\ulcorner\phi\urcorner$, and we conclude that MGT$\models_{woq}$ $\blacklozenge$ $T\ulcorner\phi\urcorner$, as desired. □

## 9.7 DISCUSSION

Objection: *You do not allow the tense operators to occur within the scope of 'T'. Your modal logic is a meta-theory in disguise. Therefore, you have failed to respond to the ghost challenge.*

The language of truth $\mathcal{L}_{ta}$ was extended by modal operators '$\square$' and '$\blacksquare$'. In this language, I have argued, we can express grounded truth-in-$\mathcal{L}_{ta}$. My opponent now asks for grounded truth of sentences containing the new modal operators themselves, that is, she asks for truth in the extended language. I admit that this is an interesting question. However, it was not my goal in this chapter; and it is not answered easily. The reason is that we do not even have a concept of semantic groundedness for the extended language. Presumably, Kripke's construction would have to be enriched by a jump operator working on the modal vocabulary. Perhaps recent work on a grounded approach to modal predicates may provide a way of doing this [Halbach and Welch, 2009; Stern, 2012]. Certainly, however, exploring this route goes beyond the scope of the present study.

We need to distinguish between two projects. My present goal is to allow an extensional theory of truth grounded in arithmetic to express groundedness. Another goal is to develop a grounded theory of truth in the modalized language $\mathcal{L}_{tam}$. This I did not attempt to do but hope to achieve elsewhere. Here, I focused on and carried out the first step of that larger project, to take the received and well understood concept of semantic groundedness for the language of truth, and show how it is expressed by a tensed theory of truth.

Objection: *As you have reminded us yourself (corollary 1), the logic of well-ordered time is not axiomatizable. Therefore, your theory MGT is based on the logic of linear time only. This means that the theory itself cannot distinguish between the intended well-founded models and non-intended because ill-founded structures. Only well-founded models, however, you have shown to be isomorphic to the Kripke stages. Hence, your key result, proposition 19, is not available to us in our own language. Therefore, the ghost challenge still stands.*

Again, we must distinguish two projects. On the one hand, one may attempt to give a formal system by which to *compute* whether or not a given sentence is grounded. However, in view of the fact that the set of grounded sentences is not computable, this is a hopeless project. I certainly did not attempt to do this.

On the other hand, we may confine ourselves to providing a means to express groundedness without semantic ascent. This has been the goal of the present chapter.

Let me elaborate on the sense in which MGT *expresses* groundedness. By way of analogy, consider how propositional tense logic expresses certain first-order concepts. For example, the propositional fragment of my logic *lq* expresses linearity in the precise sense that its axioms are valid in a frame if and only if it is a linear order (recall definition 26). This does not mean that within the propositional logic of linear time one can *define* what it is for a relation to be linear. Nonetheless, there is a good sense in which the logic *characterizes* linear frames. In fact, the power of statements in propositional modal logic to characterize a class of frames is one important reason why modal logic is interesting for logicians outside of philosophy. For example, three slogans introduce to the representative textbook Blackburn et al. [2002], the first two of which are the following [2002, pp. viii f., my emphases].

> Modal languages are simple yet expressive languages for *talking about* relational structures.

> Modal languages provide an *internal, local* perspective on relational structures.

It is in this spirit that I use the logic of linear time. The challenge I respond to is that semantic groundedness cannot be expressed within the language of truth. I distinguished this challenge from the plain fact that groundedness cannot be *defined* in its own language (p. 158). The latter follows from Tarski's theorem of the undefinability of truth. Modal logic, however, is widely used to express concepts in a setting where they cannot be defined, primarily first-order concepts in propositional modal logic.

Now, semantic groundedness is not a first-order concept. Its definition as the minimal closure of Kripke's Strong Kleene jump essentially involves second-order resources. However, the relation between my modalized first-order theory of truth and this second-order concept is analogous to the relation which a propositional modal logic bears to the class of frames that it characterizes.

Objection: *As you mentioned in the introduction (p. 160), instead of tense we may use an analogue of Boolos' stage theory to express groundedness. In fact,* prima facie *this approach is a strong competitor to yours. Stage theory is purely extensional, and also better understood than your tensed theory of truth.*

The standard, meta-theoretic definition of semantic groundedness is in terms of Kripke's least fixed point. However, as observed earlier, it is equivalent to defining groundedness in terms of the stages. To be in the fixed point is to be at some stage. Consequently, if we added the expressive means to say that sentence $\phi$ is at some stage, then it would be merely a change of notation to have a predicate of being in the least fixed point. Thus, adding to our theory of grounded truth an extensional theory of stages appears suspiciously close to adding a theory of its models. To this extent, a stage theoretic approach is a less attractive response to the ghost challenge of section 9.2.

## 9.8 CONCLUSION

The project of motivating a theory of type-free truth from the notion of groundedness faces the challenge that groundedness is a metatheoretic notion. I offered a response to this challenge. We can express the idea of groundedness in our own language using intensional means, more precisely: tense.

I presented one way of implementing this response and formulated a theory of truth based on the logic of linearly-ordered time and showed that it theory relates naturally to Kripke's semantic construction. I take this to be evidence for my proposal.

# 10 CONCLUSION

In my thesis, I have examined the concept of groundedness, as the term is used in philosophical logic. In doing so, I pursued two main objectives. Firstly, I intended to clarify and develop the concept. Secondly, my goal was to spell out its philosophical significance. Accordingly, my thesis divides into two parts. In its first three chapters I presented a general formal theory of groundedness, and applied it to paradigmatic cases from the literature, as well as put it to new use (ch. 4). The remaining five chapters motivated, developed and defended a novel account of the philosophical significance of groundedness.

I briefly summarize what has been done. Based on the primitive concept of a generator, my general theory of groundedness captures both groundedness as what can be generated (§2.3), and groundedness as having been generated (§2.4). These *prima facie* distinct intuitions were put into a common framework already by Yablo [1982]; my theory, however, not only allowed for a simplified proof of his main theorem, but also for a more finely grained concept of dependence. Further, I examined different types of generators and how their properties bear on the resulting cases of groundedness. Finally, I applied my general theory to two simple paradigms, Cantor's ordinal numbers (§ 2.6), and the pure well-founded sets (§ 6).

In chapter 3 I turned to a more complicated but philosophically very interesting case of groundedness: Kripke's semantic groundedness. I argued that the standard Kripke jump can be split into two steps, each of which corresponding to its own generator. In a first step, we generate literals of the form $T\ulcorner\phi\urcorner$ and $\neg T\ulcorner\phi\urcorner$. This *truth* generator **T** is common to all variants of Kripke's construction. They differ in the second type of generator of my *high-resolution* characterization, that correspond to closure of those literals under a specific monotone logic.

I then turned to study a case of groundedness of great interest to philosophers of mathematics, a groundedness approach to type-free class theories (chapter 4). Having laid out desiderata, I first explored the derivative approach. I translated "$x$ is in the class of the $\phi$s" as "$\phi(x)$ is true" (section IV). Through this translation, a theory of grounded truth induces a corresponding theory of grounded classes.

However, the resulting theories all proved not to give an extensional theory of classes. Therefore, I turned to developing a theory of grounded classes directly (§4.5). Building on work by Penelope Maddy and others, I proposed a new model construction that treats class identity as seriously as it does membership, and thus succeeds in characterizing a fully extensional notion of class (proposition 9). However, I also showed its theory to suffer from an impoverished schema of class comprehension (proposition 12).

Then, I moved on to the second part of my thesis, and asked for the philosophical significance of groundedness. In chapter 5 I argued that this question requires attention. Cases of groundedness, prominently the well-founded sets and Kripke's least fixed point construction, are widely considered to be significant, but no systematic account of their philosophical content had yet been put forward. By giving examples, I showed that the formal concept by itself does not ensure philosophical significance but that it requires philosophical supplementation. In the remainder of the thesis, I took first steps towards filling this need.

My starting point was a certain philosophical perspective on the well-founded sets that can be traced back to Gödel [1947] and has been discussed extensively since. According to this *iterative conception*, the cumulative hierarchy is the correct picture of the universe of sets. There are several competing ways of spelling out this thought, but they agree that at its core, the iterative conception takes a set to be constituted from its elements. I considered different attempts at analysing the relevant notion of constitution but concluded that it is best taken as a primitive relation of ontological priority. This primitive was characterized by means of examples and formal principles. I then observed that these principles are satisfied by the relation of set generation. Thus, I argued, set groundedness exemplifies the philosophical notion of ontological priority, or family of cognate notions, and to this extent, at least, is philosophically significant.

In the remainder of the thesis, I then returned to semantic groundedness, and asked for an analogous account of its significance. Chapter 7 identified a notion of priority, not of some things over others, but of some *truths* over a truth which holds in virtue of them. In the subsequent chapter, I then showed that Kripkean groundedness exemplifies this in-virtue-of relation just like set groundedness exemplifies the relation of some things consituting another. In the last chapter of my thesis, I defended this proposal against the challenge of Kripke's 'ghost of the hierarchy', that it is essentially meta-theoretic and therefore not available for the groundedness of truth in our own language.

I thus believe that my thesis has achieved both of its two main objectives. On the one hand, I have clarified the concept of groundedness. I have provided a new, systematic map of the various cases and illuminated what they have in common. On the other hand, I have taken first steps towards a philosophical account of groundedness. Having argued that the formal concepts needs philosophical supplementation, I developed a strategy to fill this need. Groundedness by some generator is philosophically significant at least to the extent that this generator tracks some philosophical notion of priority. In particular, I developed a novel account of semantic groundedness, inspired by the philosophy of set theory: an *iterative conception* of truth which uses the in-virtue-of relation of contemporary metaphysics just like the iterative conception of sets bases on a notion of ontological constitution.

I close by outlining routes of future research. The main next step will be to test my proposal further. This can be done in various philosophically fruitful ways. For one, I intend to develop an analogous account of the concept of grounded classes, or concept-extensions, thus supplementing the logical study of chapter 4 with a philosophical one. A case can be made that concept-extensions should not be thought of as constituted from their members, at least not in the same sense as a set is constituted from its elements. It can further be argued that what matters instead is whether something falls under the defining concept, and if so, in virtue of what this is the case. This line of thought suggests a connection to my proposal from chapter 8 and would allow me to test it for fruitfulness and applicability. The goal is an *iterative conception* of classes, analogous to the iterative conception of sets and my iterative conception of truth.

For another, I would like to understand better why groundedness, as in the case of sets and truth, is widely considered a satisfactory response to paradox. I am confident that my philosophical account of these cases provides a starting point towards an answer. In particular, I plan to connect with a general analysis of paradox along the lines of Russell [1908]; Priest [1994]. On this basis, paradox may be viewed as a consequence of enforcing circular priority relations. Thus, assuming that some set is an element of itself amounts by the account of chapter 6 to the assumption that something partly constitutes itself. Or, assuming the truth or falsity of a liar sentence amounts by the account of chapter 8 to the claim that some truths holds in virtue of itself. I hope that this priority-driven perspective on paradox allows for a novel argument that the groundedness approach is superior to its hierarchical competitors, such as type theory respectively Tarski's stratified theory of truth.

In addition to these tests of my philosophical proposal, I also plan to develop the technical contributions of the present thesis. For one, other applications of my general concept of groundedness ought to be examined. I am confident that certain interesting theories from the philosphical literature would be illuminated if put into my general framework [Fine, 2005; Linnebo, 2006]. Also, entirely novel applications seem possible, such as to a theory of intensionality informed by the Russell-Myhill paradox. For another, I hope to soon generalize the modal machinery of chapter 9 and be able to present for every instance of the general concept a corresponding modal theory.

# BIBLIOGRAPHY


Peter Aczel. An introduction to inductive definitions. In Jon Barwise, editor, *Handbook of mathematical logic*, pages 739–782. Amsterdam, 1977. (Cited on pages 22, 28, 30, and 35.)

Peter Aczel. Frege structures and the notions of proposition, truth and set. In Jon Barwise, Howard Jerome Keisler, and Kenneth Kunen, editors, *The Kleene Symposium*, pages 31–59. North-Holland, 1980. (Cited on page 63.)

Aristotle. *Metaphysics Book Θ*. Clarendon Press, 2006. translated with an introduction and commentary by Stephen Makin. (Cited on page 120.)

Jon Barwise. *Admissible Sets and Structures*. Perspectives in Mathematical Logic. Springer, Berlin Heidelberg New York, 1975. (Cited on page 28.)

Roderick Batchelor. Grounds and consequences. *Grazer Philosophische Studien*, 80(1):65–77, 2010. (Cited on pages 134, 142, and 143.)

JC Beall. Prolegomenon to Future Revenge. In JC Beall, editor, *Revenge of the Liar: New Essays on the Paradox*, pages 1 – 30. Oxford University Press, Oxford, 2007. (Cited on page 158.)

Jan Berg. *Bolzano's Logic*. Almqvist & Wiksell, Stockholm, 1962. (Cited on page 121.)

P. Blackburn, M. De Rijke, and Y. Venema. *Modal logic*, volume 53. Cambridge University Press, 2002. (Cited on page 183.)

Stephen Blamey. Partial Logic. In *Handbook of Philosophical Logic*, volume 5, pages 262 – 353. D.Reidel, second edition edition, 2002. (Cited on page 46.)

George Boolos. The Iterative Conception of a Set. *The Journal of Philosophy*, 68:215–232, 1971. (Cited on pages 95, 112, and 160.)

George Boolos. Iteration again. In *Logic, Logic and Logic*. Harvard University Press, 1998. (Cited on page 108.)

Ross T. Brady. The consistency of the axioms of abstraction and extensionality in a three-valued logic. *Notre Dame Journal of Formal Logic*, 12(4):447–453, 1971. (Cited on pages 65 and 85.)

John P. Burgess. The Truth is Never Simple. *The Journal of Symbolic Logic*, 51(3):663 – 681, 1986. (Cited on pages 56 and 57.)

John P. Burgess. E Pluribus Unum: Plural Logic and Set Theory. *Philosophia Mathematica*, 12:193 – 221, 2004. (Cited on page 22.)

John P. Burgess. Friedman and the Axiomatization of Kripke' Theory of Truth. Unpublished manuscript, delivered at a conference in honour of the 60th birthday of Harvey Friedman at the Ohio State University, 14-17 May 2009, 2009. (Cited on page 70.)

Andrea Cantini. A Theory of Truth Formally Equivalent to id1. *Journal of Symbolic Logic*, 55:244 – 258, 1990. (Cited on pages 56 and 80.)

Andrea Cantini. *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*, volume 135 of *Studies in Logic and the Foundations of Mathematics*. Elsevier, 1996. (Cited on pages 63, 65, and 72.)

Georg Cantor. Foundations of a general theory of manifolds: a mathematico-philosophical investigation into the theory of the infinite. In William B. Ewald, editor, *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, volume 2, pages 878 – 919. Clarendon Press, Oxford, 1883. (Cited on page 104.)

Georg Cantor. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Springer, 1932. edited by Ernst Zermelo. (Cited on page 37.)

John H. Conway. *On Numbers and Games*. Academic Press, Natick, MA, 2nd edition, 2001. (Cited on page 96.)

Fabrice Correia. Husserl on foundation. *Dialectica*, 58(3):349–367, 2004. (Cited on page 115.)

Fabrice Correia. Ontological Dependence. *Philosophy Compass*, 3, 2008. (Cited on page 115.)

Fabrice Correia. Grounding and Truth-functions. *Logique et Analyse*, 116, 2011. (Cited on page 142.)

Fabrice Correia. Logical grounds. *Review of Symbolic Logic*, 2013. forthcoming. (Cited on pages 132, 142, 144, 145, 152, and 153.)

Fabrice Correia and Benjamin Schnieder. Grounding: An opinionated introduction. In Fabrice Correia and Benjamin Schnieder, editors, *Metaphysical grounding: Understanding the structure of reality*. Cambridge University Press, Cambridge, New York, Melbourne, 2012. (Cited on page 135.)

Shamik Dasgupata. On the plurality of grounds. unpublished manuscript, 2013. URL `http://www.shamik.net/papers/dasguptaunds.pdf`. (Cited on page 142.)

Michael Dummett. The philosophical significance of godel's theorem. In *Truth and other enigmas*, pages 186–202. Duckworth, London, 1978. (Cited on page 38.)

William B. Ewald. *From Kant to Hilbert: Readings in the Foundations of Mathematics*. Clarendon Press, 1996. ISBN 0198532717. (Cited on page 37.)

Solomon Feferman. Non-extensional type-free theories of partial operations and classifications, i. In *ISILC Proof Theory Symposion*, pages 73–118. Springer, 1975a. (Cited on page 63.)

Solomon Feferman. A language and axioms for explicit mathematics. In J. N. Crossley, editor, *Algebra and logic*, volume 450 of *Lecture Notes in Mathematics*, pages 87–139. Springer, 1975b. (Cited on pages 63, 64, and 68.)

Solomon Feferman. Constructive theories of functions and classes. In *Logic Colloqium '78*, volume 97 of *Studies in Logic and the Foundations of Mathematics*, pages 159–224. Elsevier, 1979. (Cited on page 64.)

Solomon Feferman. Reflecting on Incompleteness. *The Journal of Symbolic Logic*, 56:1–49, 1991. (Cited on pages 67 and 72.)

Solomon Feferman. Axioms for determinateness and truth. *Review of Symbolic Logic*, 1(1):204–217, 2008. (Cited on page 54.)

Jose Ferreiros. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*. Birkhäuser Verlag, Basel - Boston - Berlin, 2007 edition, 1999. (Cited on page 39.)

Kit Fine. The study of ontology. *Noûs*, 25(3):263–294, 1991. (Cited on page 23.)

Kit Fine. Part-whole. In Barry Smith and David Woodruff Smith, editors, *The Cambridge Companion to Husserl*. cup, Cambridge, 1995a. (Cited on page 115.)

Kit Fine. Ontological Dependence. *Proceedings of the Aristotelian Society*, 95, 1995b. (Cited on pages 115 and 116.)

Kit Fine. The question of realism. *Philosophers' Imprint*, 1(1):1–30, 2001. (Cited on pages 120, 127, 129, 134, and 142.)

Kit Fine. Class and Membership. *The Journal of Philosophy*, 102(11), November 2005. (Cited on page 188.)

Kit Fine. Some Puzzles of Ground. *Notre Dame Journal of Formal Logic*, 51(1):97–118, 2010. (Cited on pages 49, 132, 139, 140, 141, 152, 153, and 154.)

Kit Fine. The pure logic of ground. *The Review of Symbolic Logic*, 5(1): 1–25, 2012a. (Cited on pages 124, 129, 145, and 149.)

Kit Fine. Guide to ground. In Fabrice Correia and Benjamin Schnieder, editors, *Metaphysical grounding: Understanding the structure of reality*, pages 37–80. Cambridge University Press, Cambridge, New York, Melbourne, 2012b. (Cited on pages 19, 27, 120, 141, 142, 145, 146, 149, and 151.)

Martin Fischer, Volker Halbach, Jönne Kriener, and Johannes Stern. Axiomatizing semantic theories of truth. manuscript, 2014. (Cited on pages 57 and 173.)

Thomas Forster. The iterative conception of set. *Review of Symbolic Logic*, 1(1):97–110, 2008. (Cited on pages 17, 22, 27, 35, 39, 41, 92, 93, 94, 95, 96, 97, and 98.)

A.A. Fraenkel, Y. Bar-Hillel, and A. Levy. *Foundations of set theory*, volume 67. North Holland, 1973. (Cited on page 64.)

Harvey Friedman and Michael Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33:1–21, 1987. (Cited on pages 100, 101, 102, and 103.)

Kentaro Fujimoto. Relative truth definability of axiomatic truth theories. *Bulletin of Symbolic Logic*, 16(3):305–344, 2010. (Cited on page 54.)

Kentaro Fujimoto. Classes and truths in set theory. *Annals of Pure and Applied Logic*, 163:1484–1523, 2012. (Cited on pages 72 and 79.)

J. W. Garson. Quantification in modal logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 1st edition, volume2: Extensions of Classical Logic*. Reidel, Dordrecht, 1984. (Cited on page 162.)

P. C. Gilmore. The consistency of partial set theory without extensionality. In T. Jech, editor, *Axiomatic Set Theory II*, Proceedings of Symposia in Pure Mathematics. American Mathematical Society, Providence, Rhode Island, 1974. (Cited on page 73.)

Georges Gonthier. Formal proof—the four-color theorem. *Notices of the American Mathematical Society*, 55(11):1382–1393, 2008. (Cited on page 92.)

Kurt Gödel. What is cantor's continuum problem? In Solomon Feferman, John W. Dawson Jr., Warren Goldfarb, Charles Parsons, and Robert M. Solovay, editors, *Collected Works II*, pages 154 – 187. Oxford University Press, Oxford, 1947. (Cited on pages 39 and 187.)

Volker Halbach. *Axiomatische Wahrheitstheorien*. Akademie Verlag, 1996. (Cited on page 167.)

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011. (Cited on pages 50, 53, 65, 70, 82, 103, 166, 170, 171, 174, and 181.)

Volker Halbach and Philip Welch. Necessities and necessary truths: A prolegomenon to the use of modal logic in the analysis of intensional notions. *Mind*, 118(469), 2009. (Cited on page 182.)

Allen Hazen. Davis's formulation of kripke's theory of truth: A correction. *Journal of Philosophical Logic*, 10:309–311, 1981. (Cited on page 31.)

Richard Heck. The development of arithmetic in frege's grundgesetze der arithmetik. *Journal of Symbolic Logic*, 58(2):579–600, 1993. (Cited on page 92.)

Richard G. Heck. The Consistency of Predicative Fragments of Frege's Grundgesetze der Arithmetik. *History and Philosophy of Logic*, 17: 209–220, 1996. (Cited on page 81.)

Hans G. Herzberger. Paradoxes of Grounding in Semantics. *Journal of Philosophy*, 67:145 – 167, 1970. (Cited on pages 22 and 44.)

R Hinnion. Le paradoxe de russell dans des versions positives de la théorie naïve des ensembles. *CR Acad. Sci. Paris, Sér. I Math*, 304: 307–310, 1987. (Cited on page 73.)

Roland Hinnion and Thierry Libert. Positive abstraction and extensionality. *Journal of Symbolic Logic*, 68(3):828–836, 2003. (Cited on pages 65 and 73.)

Leon Horsten. *The Tarskian Turn: Deflationism and Axiomatic Truth*. MIT Press, Cambridge, MA, 2011. (Cited on page 45.)

Edmund Husserl. *Logical Investigations*, volume 2. Routledge, London and New York, 2001. translated by J.N. Findlay from the Second German edition. (Cited on page 115.)

Luca Incurvati. How to be a minimalist about sets. *Philosophical Studies*, pages 1–19, 2012. (Cited on page 108.)

Ingasi Jané. Idealist and Realist Elements in Cantor's Approach to Set Theory. *Philosophia Mathematica*, 18:193 – 226, 2010. (Cited on page 37.)

Gerhard Jäger, Reinhard Kahle, and Thomas Studer. Universes in explicit mathematics. *Annals of Pure and Applied Logic*, 109:141–162, 2001. (Cited on page 64.)

Jeffrey Ketland. Identity and Indiscernibility. *The Review of Symbolic Logic*, 4(02):171–185, 2011. (Cited on page 85.)

Saul Kripke. Outline of a Theory of Truth. *The Journal of Philosophy*, 72(19):690 – 716, November 1975. Seventy-Second Annual Meeting Americal Philosophical Association. (Cited on pages 22, 44, 45, 47, and 65.)

Saul A. Kripke. *Naming and Necessity*. Harvard University Press, Cambridge, Massacusetts, 1980. (Cited on page 114.)

Hannes Leitgeb. What Truth Depends On. *Journal of Philosophical Logic*, 35:155–192, 2005. (Cited on page 22.)

Stephan Leuenberger. Grounding and necessity. 57(2), April 2014. (Cited on page 142.)

David Lewis. *On the plurality of Worlds*. Blackwell Publisher, Malden, Oxford, 1986. (Cited on page 114.)

David Lewis. *Parts of Classes*. Basil Blackwell, 1991. (Cited on page 115.)

Øystein Linnebo. Sets, Properties and Unrestricted Quantification. In Agustin Rayo and Gabriel Uzquiano, editors, *Absolute Generality*, pages 149 – 178. Clarendon Press, Oxford, 2006. (Cited on pages 64 and 188.)

Øystein Linnebo. The Potential Hierarchy of Sets. *The Review of Symbolic Logic*, 6:205–228, 6 2013. ISSN 1755-0211. doi: 10.1017/S1755020313000014. URL http://journals.cambridge.org/article_S1755020313000014. (Cited on page 113.)

Penelope Maddy. Proper Classes. *Journal of Symbolic Logic*, 48(1):113–139, Mar 1983. (Cited on pages 63, 65, 78, and 79.)

Penelope Maddy. Believing the Axioms I. *The Journal of Symbolic Logic*, 53(2):481–511, June 1988. (Cited on page 64.)

Penelope Maddy. A theory of sets and classes. In Gila Sher and Richard Tieszen, editors, *Between Logic and Intuition: Essays in Honor of Charles Parsons*, pages 299–316. Cambridge University Press, 2000. (Cited on pages 63, 65, and 78.)

Tim Maudlin. *Truth and Paradox: Solving the Riddles*. Clarendon Press, Oxford, 2004. (Cited on pages 22 and 44.)

Timothy McCarthy. Ungroundedness in Classical Languages. *Journal of Philosophical Logic*, 17(1):61–74, 1988. (Cited on page 22.)

Vann McGee. *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*. Hackett, Indianapolis, Cambridge, 1991. (Cited on pages 29, 44, and 45.)

Yiannis N. Moschovakis. *Elementary Induction on Abstract Structures*. North-Holland, 1974. (Cited on pages 22 and 28.)

Urs Oswald. *Fragmente von "New Foundations" und Typentheorie*. PhD thesis, ETH Zürich, 1976. (Cited on page 93.)

Charles Parsons. Sets and classes. *Noûs*, 8(1):pp. 1–12, 1974. (Cited on page 64.)

Charles Parsons. What is the Iterative Conception of Set? In Robert E. Butts and Jaakko Hintikka, editors, *Logic, Foundations of Mathematics and Computability Theory*, volume 9 of *University of Western Ontario Series in Philosophy of Science*, pages 335–367. Reidel, Dordrecht, 1977. (Cited on pages 63, 108, 109, and 110.)

Charles Parsons. *Mathematics in Philosophy: Selected Essays*, chapter Sets and Modality. Cornell University Press, Ithaca, New York, 1983. (Cited on pages 110, 111, and 112.)

Michael Potter. *Set Theory and its Philosophy: A Critical Introduction*. Oxford University Press, Oxford, 2004. (Cited on pages 95, 108, 109, 112, and 160.)

Graham Priest. The structure of the paradoxes of self-reference. *Mind*, 103(409):25–34, 1994. (Cited on page 188.)

Graham Priest. *An Introduction to Non-Classical Logic: From* If *to* Is. Cambridge University Press, Cambridge, second edition, 2008. (Cited on page 164.)

W. V. Quine. *The Ways of Paradox, and Other Essays*. Harvard University Press, 1976. (Cited on page 92.)

Agustin Rayo and Gabriel Uzquiano. *Absolute Generality*. Clarendon Press, Oxford, 2006. (Cited on page 95.)

Gideon Rosen. Metaphysical Dependence: Grounding and Reduction. In Bob Hale and Aviv Hoffman, editors, *Modality: Metaphysics, Logic and Epistemology*. Oxford University Press, Oxford, 2010. (Cited on page 142.)

Antje Rumberg. Bolzano's concept of grounding (*Abfolge*) against the background of normal proofs. *The Review of Symbolic Logic*, 2013. (Cited on page 122.)

Antje Rumberg and Stefan Roski. Bolzano on the hierarchy of truths. unpublished manuscript, 2012. (Cited on page 143.)

Bertrand Russell. Mathematical Logic as Based on a Theory of Types. *American Mathematical Journal*, 30:222 – 262, 1908. (Cited on page 188.)

Jonathan Schaffer. On What Grounds What. In David Chalmers, David Manley, and Ryan Wasserman, editors, *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press, Oxford, 2009. (Cited on pages 115, 116, and 120.)

Benjamin Schnieder. A Logic for Because. *The Review Of Symbolic Logic*, 4(3), September 2011. (Cited on pages 135, 139, 140, and 142.)

Stewart Shapiro. *Foundations without Foundationalism: a case for second-order logic*. Oxford Logic Guides. Clarendon Press, 1991. (Cited on page 25.)

Stewart Shapiro and Crispin Wright. All things indefinitely extensible. In Agustin Rayo and Gabriel Uzquiano, editors, *Absolute Generality*, pages 255–304. Clarendon Press, Oxford, 2006. (Cited on page 38.)

Joseph R. Shoenfield. Axioms of set theory. In Jon Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *of Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1977. (Cited on pages 108 and 109.)

Theodore Sider. *Writing the Book of the World*. Oxford University Press, 2011. (Cited on page 116.)

Peter Simons. Bolzano on collections. In Mark Siebel Wolfgang Künne and Mark Textor, editors, *Bolzano and Analytic Philosophy*, pages 87 – 108. Rodopi, Amsterdam, 1997. (Cited on page 122.)

Eliott Sober. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. University of Chicago Press, 1984. (Cited on page 92.)

Johannes Stern. *Toward Predicate Approaches to Modality*. PhD thesis, University of Geneva, 2012. (Cited on page 182.)

James P. Studd. The iterative conception of set: A (bi-)modal axiomatisation. *Journal of Philosophical Logic*, 2013. (Cited on page 113.)

Alfred Tarski. *The Concept of Truth in Formalized Languages*, pages 152 – 278. Clarendon Press, Oxford, 1956.

Gabriel Uzquiano. Plural Quantification and Classes. *Philosophia Mathematica*, 11:67–81, 2003. (Cited on page 64.)

Albert Visser. Semantics and the Liar Paradox. In Dov Gabbay and Franz Günther, editors, *Handbook of Philosophical Logic*. Kluwer, 1983. (Cited on pages 44, 50, and 161.)

Hao Wang. Large sets. In Robert E. Butts and Jaakko Hintikka, editors, *Logic, Foundations of Mathematics and Computability Theory*, volume 9 of *University of Western Ontario Series in Philosophy of Science*, pages 309 – 335. Reidel, Dordrecht, 1977. (Cited on page 109.)

Hermann Weyl. *Philosophy of Mathematics and Natural Science*. Princeton University Press, Princeton, 1949. (Cited on page 109.)

Steve Yablo. Grounding, Dependence and Paradox. *Journal of Philosophical Logic*, 11(1):117–137, 1982. (Cited on pages 22, 25, 42, 47, 54, and 186.)